

**LAPORAN PRAKTIKUM**  
**DATA WAREHOUSING DAN DATA MINING**

**PERTEMUAN 12**  
**“REGRESI LINIER SEDERHANA”**



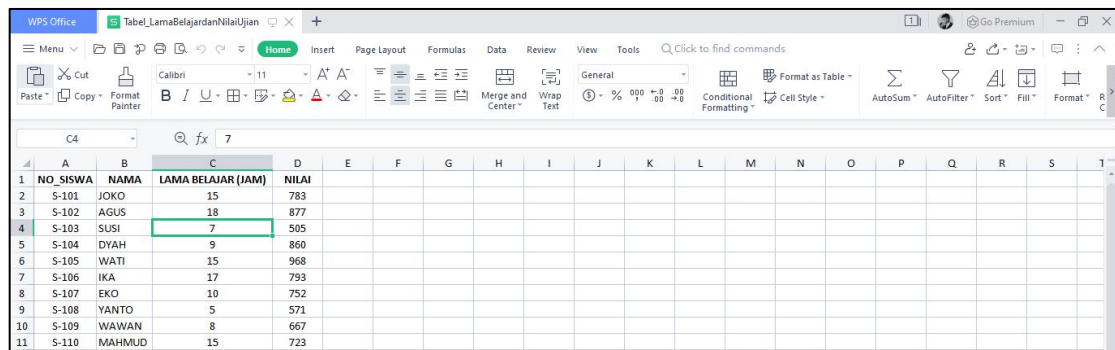
**Oleh:**

**NAMA : Daffa Putra Alwansyah**  
**NIM : L200190031**  
**KELAS : B**  
**PRODI : INFORMATIKA**

**Fakultas Komunikasi dan Informatika**  
**Universitas Muhammadiyah Surakarta**

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

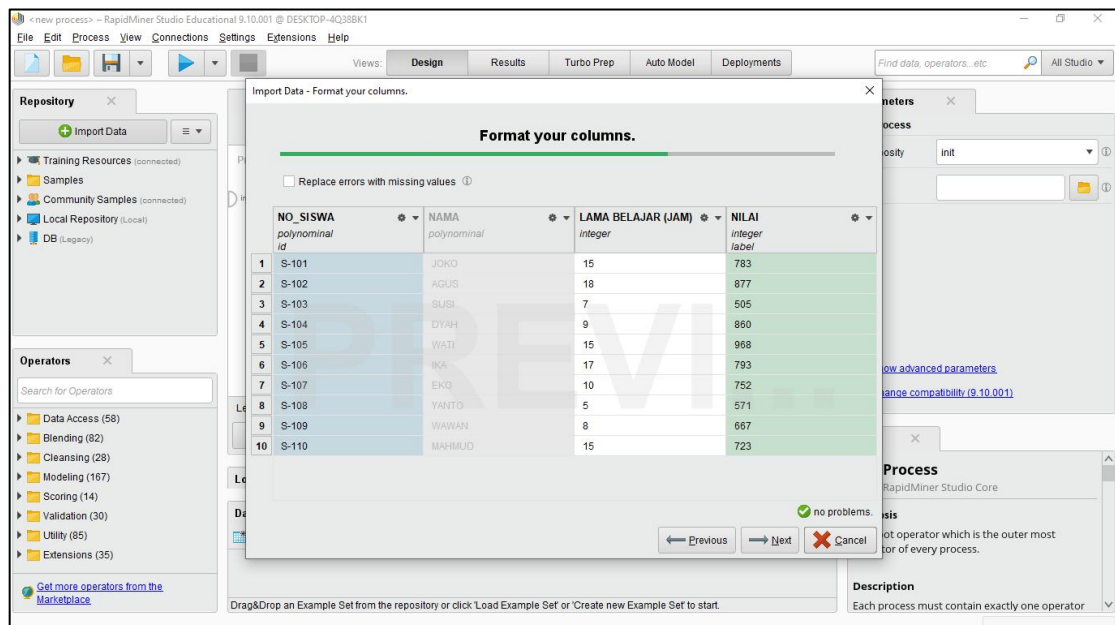
1. Membuat tabel Tabel\_LamaBelajardanNilaiUjian.xlsx, lalu masuk ke RapidMiner.



The screenshot shows a WPS Office spreadsheet with the following data:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
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																

2. Ubah tipe data menjadi: NO\_SISWA : polynomial, id | NAMA : exclude column | LAMA JAM BELAJAR : integer | NILAI : integer, label.



The screenshot shows the RapidMiner Studio interface with the 'Format your columns' dialog box open. The dialog box displays the following column configurations:

Column	Type
NO_SISWA	polynomial id
NAMA	polynomial
LAMA BELAJAR (JAM)	integer
NILAI	integer label

The dialog box also includes a 'Replace errors with missing values' checkbox and 'Previous', 'Next', and 'Cancel' buttons. The background shows the 'Repository' and 'Operators' panels.

3. Ubah menjadi Data\_LamaBelajardanNilaiUjian.



The screenshot shows the 'Name' and 'Location' fields in the RapidMiner Studio interface. The 'Name' field is set to 'Data\_LamaBelajardanNilaiUjian' and the 'Location' field is set to '//Local Repository/Data\_LamaBelajardanNilaiUjian'. The 'Previous', 'Finish', and 'Cancel' buttons are visible at the bottom.

- Masukan Data dan operator *Linear Regression*, ubah min tolerance 0.05, lalu run.



- Berikut hasil proses regresi linear :

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

Result History | LinearRegression (Linear Regression) | ExampleSet (/Local Repository/Data\_LamaBelajardanNilaiUjian)

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	****

- Text View (mencari model regresi).

Result History | LinearRegression (Linear Regression) | ExampleSet (/Local Repository/Data\_LamaBelajardanNilaiUjian)

**LinearRegression**

21.608 \* LAMA BELAJAR (JAM)  
+ 492.769

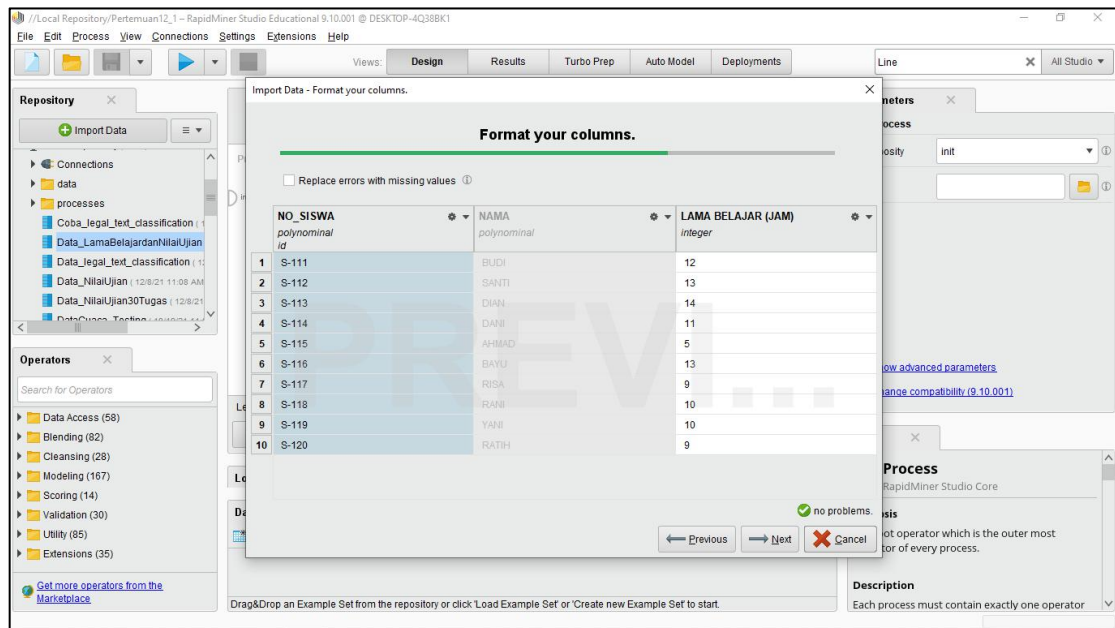
## 12.4.2 Mencari Nilai t dan Model Regresi Linier menggunakan RapidMiner

- Membuat Tabel data siswa dengan nama Tabel\_PrediksiNilaiUjian.xls, lalu masuk ke RapidMiner.

WPS Office | Tabel\_PrediksiNilaiUjian.xls

NO SISWA	NAMA	LAMA BELAJAR (JAM)
1	S-111	BUDI
2	S-112	SANTI
3	S-113	DIAN
4	S-114	DANI
5	S-115	AHMAD
6	S-116	BAYU
7	S-117	RISA
8	S-118	RANI
9	S-119	YANI
10	S-120	RATIH

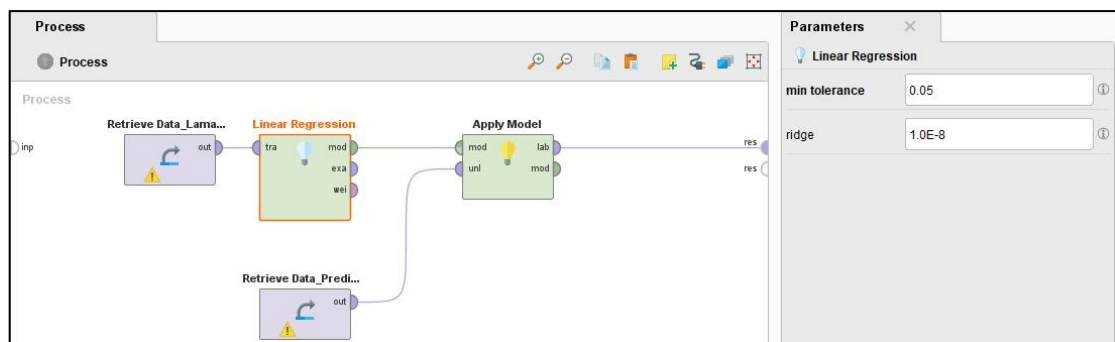
- Ubah tipe data menjadi: NO\_SISWA : polynomial, id | NAMA : exclude column  
| LAMA JAM BELAJAR : integer.



- Beri nama menjadi Data\_PrediksiNilaiUjian.

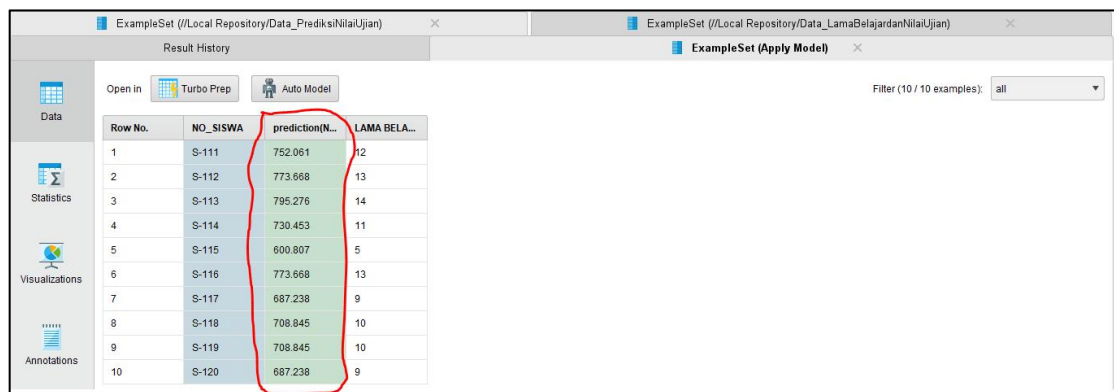


- Masukan kedua file tadi, lalu tambahkan Linear Regression dan Apply Model, hubungkan port sesuai di modul, lalu run.



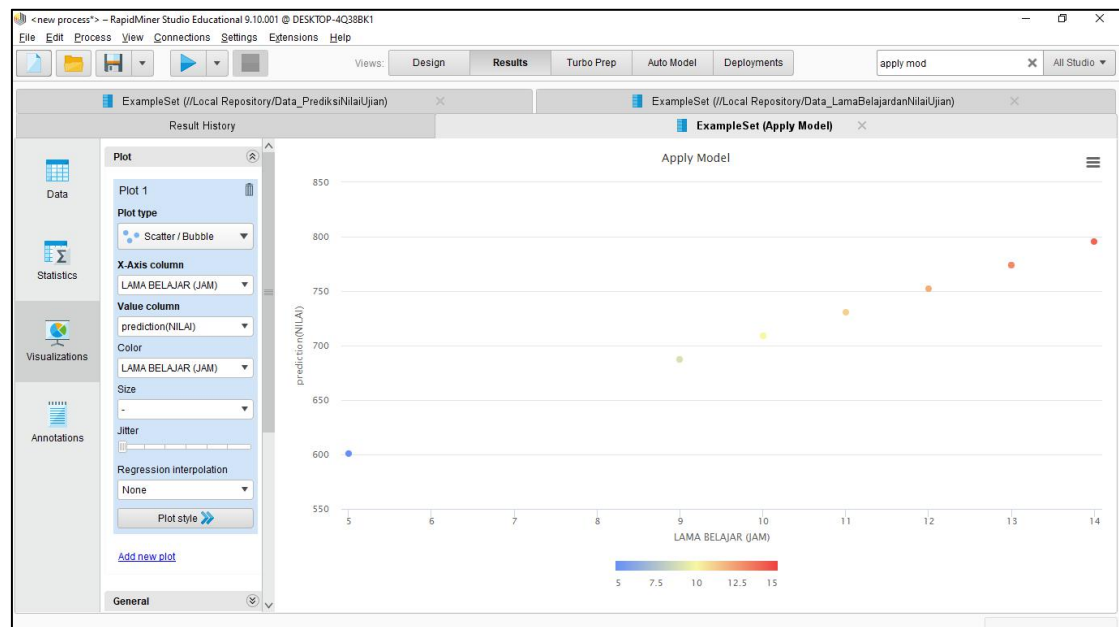
5. Berikut hasil proses prediksi terhadap data testing menggunakan regresi linear :

a) Data View (hasil prediksi nilai ujian)



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

b) Charts View (Scatter Plot)



### 12.4.3 Pembuktian Model Regresi

1. Pada tahap ini akan melakukan pembuktian hasil prediksi dengan memasukan sel E3 “ $=(21,608 \times C3)+492,769$ ” lalu tarik hingga E12.

The screenshot shows a WPS Office spreadsheet titled 'Pembuktian.xlsx'. The formula bar displays  $=(21,608 \times C3)+492,769$ . The spreadsheet contains two tables: 'Tabel' and 'Model Regresi'.

	A	B	C	D	E
1	NO_SISWA	NAMA	LAMA BELAJAR (JAM)	Prediction (NILAI)	Prediction (NILAI)
2				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

### 12.5 TUGAS

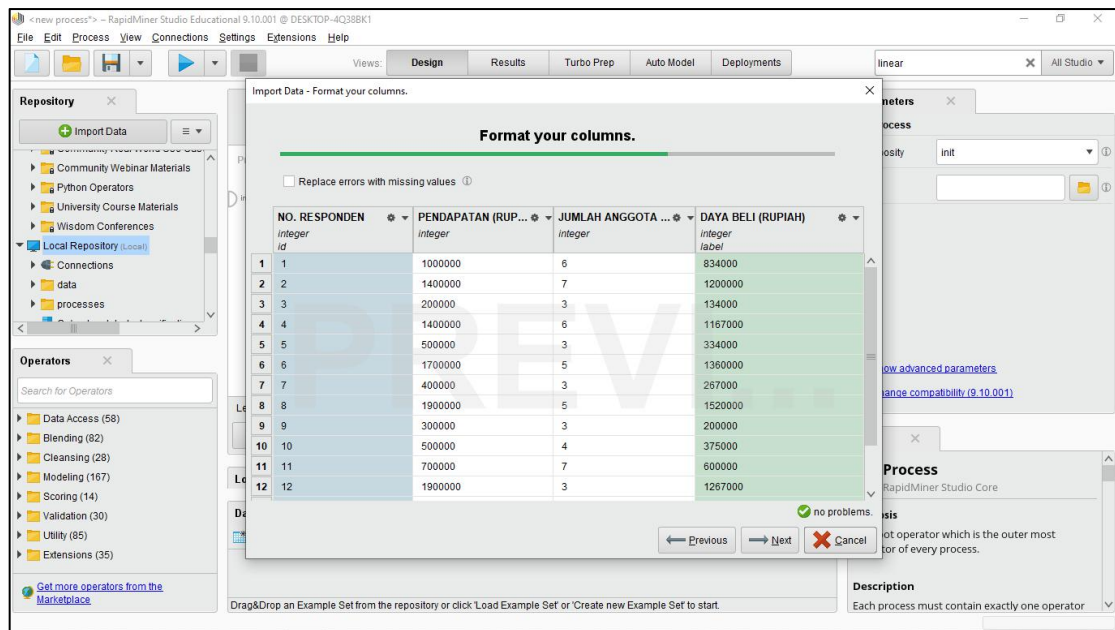
1. Membuat tabel dengan nama Tabel\_HasilSurvei, lalu ke RapidMiner.

The screenshot shows a WPS Office spreadsheet titled 'Tabel\_HasilSurvei.xlsx'. The formula bar displays 'NO. RESPONDEN'. The spreadsheet contains a table with survey data.

	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	1.125.000
16	15	1.300.000	7	1.115.000



2. Ubah sesuai role yang ada.



6. Untuk **Linear Sederhana**, dengan min tolerance 0.05/5%, lalu run.



7. Berikut hasil proses regresi linear :

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

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

Disini bisa dilihat bahwa nilai t-hitung (t-stat) ialah **35.037 dan 6.103**.

- Apabila t-hitung = 35,037 dan t-table = 2,131, maka  $35,037 > 2,131$  dengan nilai toleransi 5% (0,05). Sehingga dapat dikatakan bahwa PENDAPATAN (X1) mempengaruhi secara signifikan terhadap Daya Beli (Y).
- Apabila t-hitung = 6,103 dan t-table = 2,131, maka  $6,103 > 2,131$  dengan nilai toleransi 5% (0,05). Sehingga dapat dikatakan bahwa JUMLAH ANGGOTA (X2) mempengaruhi secara signifikan terhadap Daya Beli (Y).

b) Text View (mencari model regresi).

**LinearRegression**

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

Model Regresi yang terbentuk ialah :

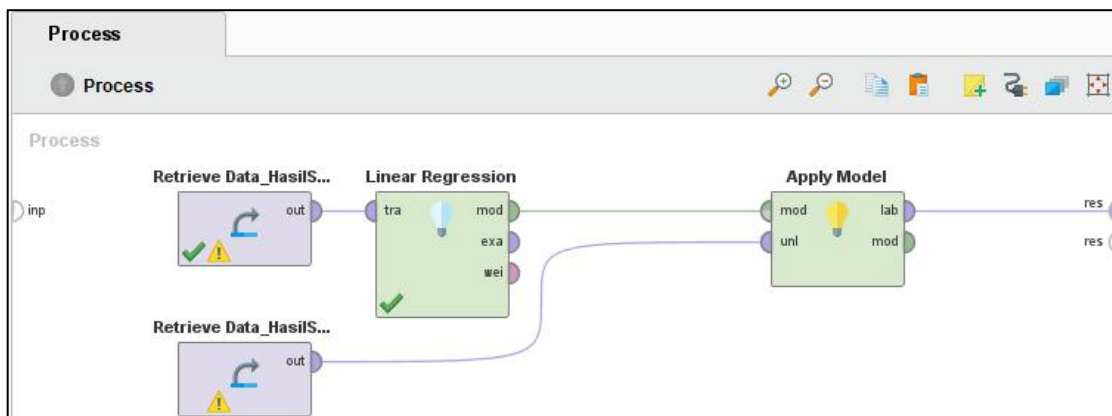
i.  $Y (\text{Daya Beli}) = 0.739 * \text{PENDAPATAN (Rupiah)} + 47807.624 * \text{JUMLAH ANGGOTA KELUARGA} - 180222.487$  atau

ii.  $Y (\text{Daya Beli}) = 0.739 X_1 + 47807.624 X_2 - 180222.487$

8. Model Persamaan Regresi Linear

	A	B	C	D	E	F	G
	NO. RESPONDEN	PENDAPATAN (RUPIAH)	JUMLAH ANGGOTA KELUARGA	PREDICTION (NILAI) TABEL	PREDICTION (NILAI) MODEL REGRESI		
1	1	900.000	5	7.239.332.625.273.290	723.916		
2	2	800.000	3	5.544.160.562.396.610	554.400		
3	3	500.000	2	2.849.025.555.614.690	284.893		
4	4	1.900.000	6	15.107.604.763.017.200	1.510.723		
5	5	600.000	2	35.880.451.457.521.200	358.793		
6	6	800.000	5	6.500.313.035.135.860	650.016		
7	7	1.000.000	6	8.456.428.451.780.350	845.623		
8	8	1.100.000	4	8.239.295.569.178.530	823.908		
9	9	1.000.000	4	7.500.275.979.041.100	750.008		
10	10	500.000	3	332.710.179.198.432	332.700		

9. Melakukan prediksi Daya Beli (Y), gunakan data testing untuk melakukan prediksi.





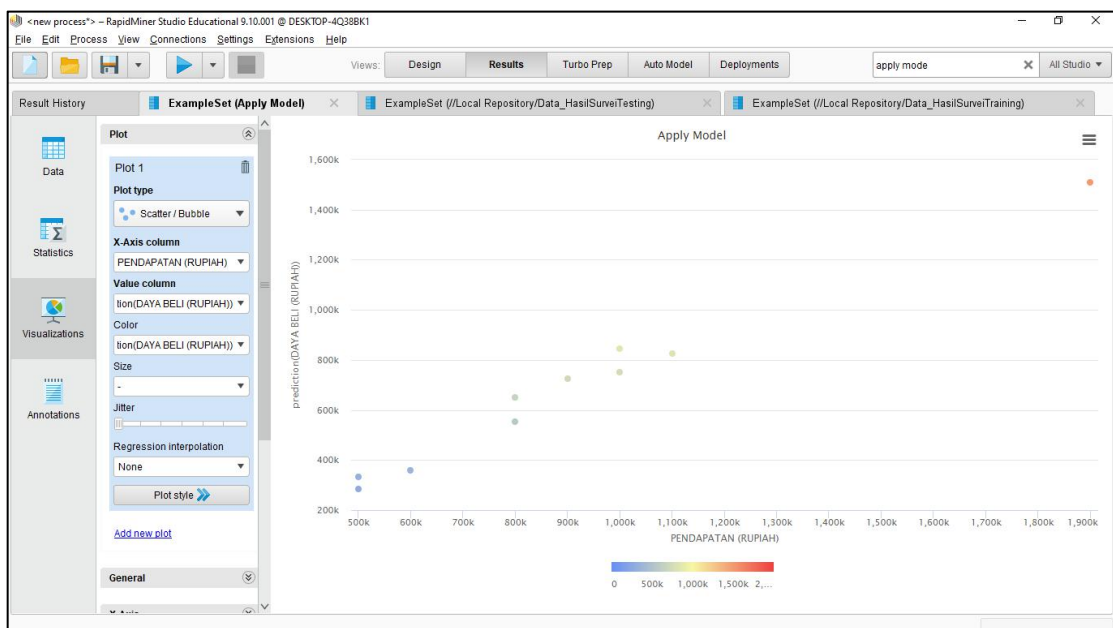
Result History					
ExampleSet (Apply Model)					
Open in <span>Turbo Prep</span> <span>Auto Model</span>					
	Row No.	NO. RESPON...	JUMLAH AN...	prediction(D...	PENDAPATA...
	1	1	5	723933.263	900000
	2	2	3	554416.056	800000
	3	3	2	284902.556	500000
	4	4	6	1510760.476	1900000
	5	5	2	358804.515	600000
	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

10. Menggambar pola sebaran data menggunakan Plot View (Scatter):

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))

