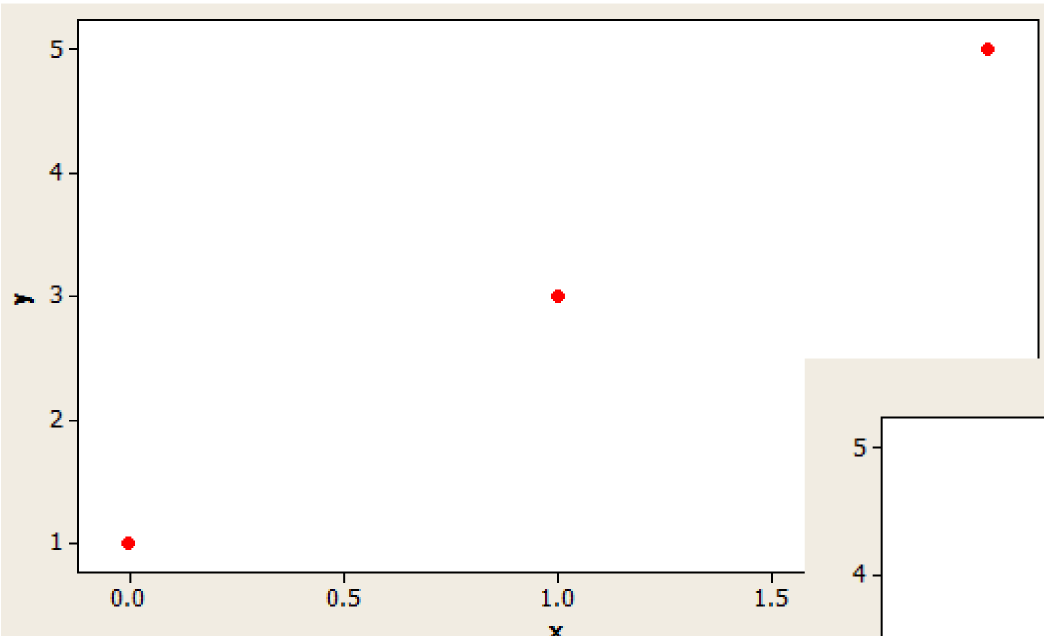
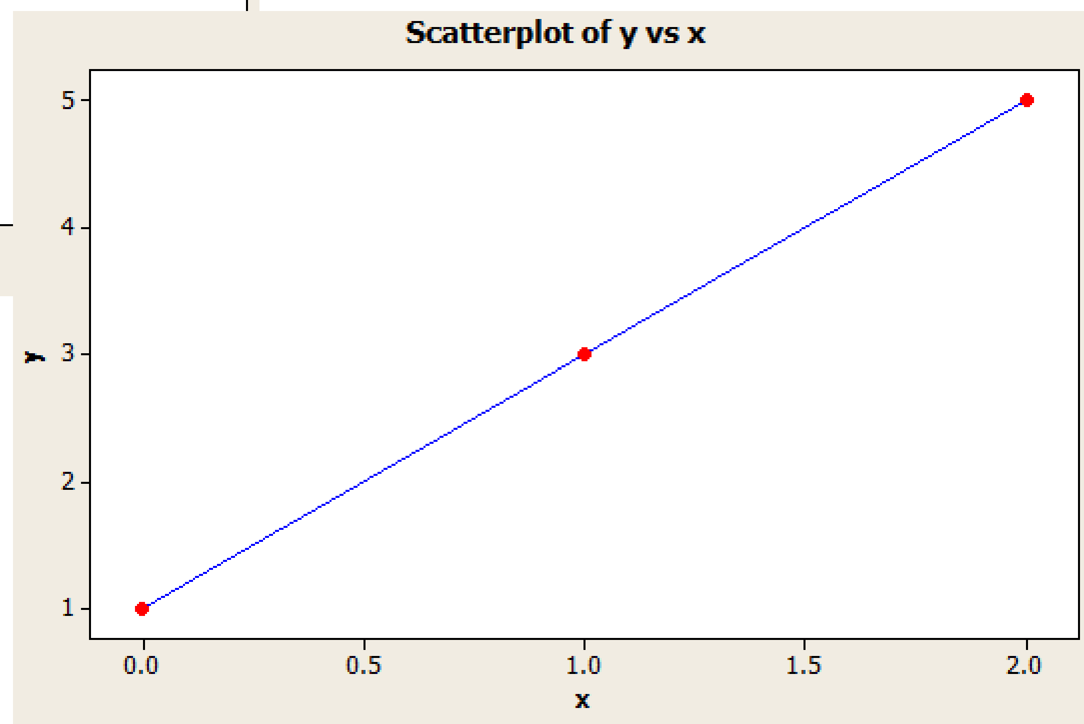
CHAPTER 3

ANALISIS REGRESI

No 1

Gambar grafik fungsi $y = 2x + 1$ dengan x=0,1,2.

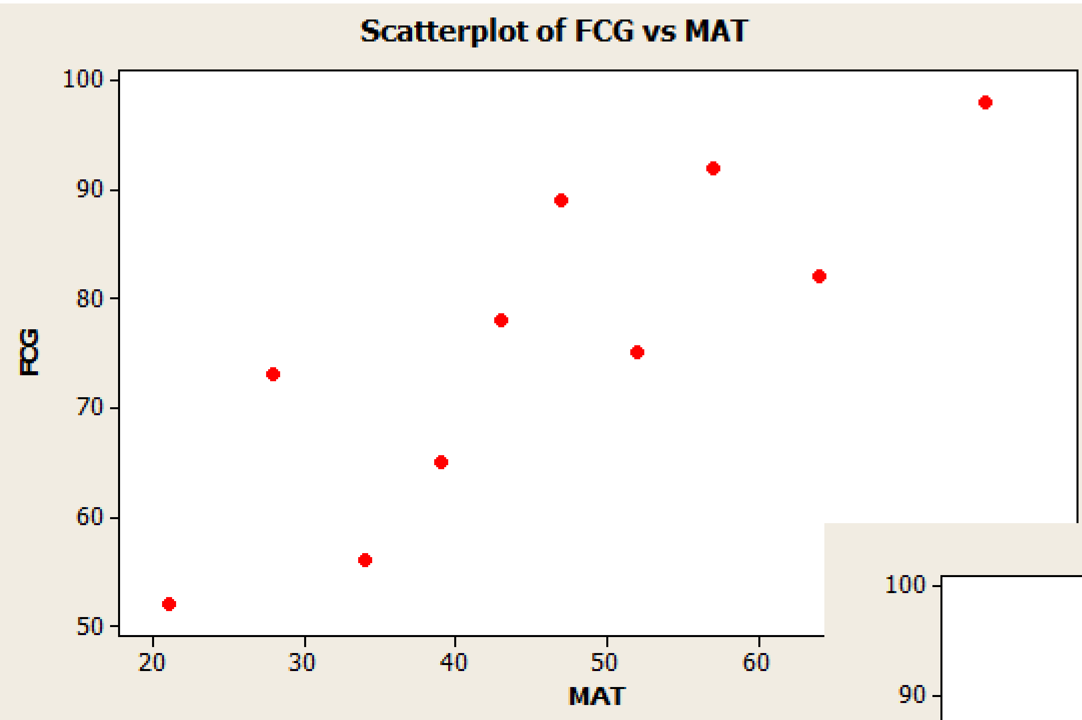


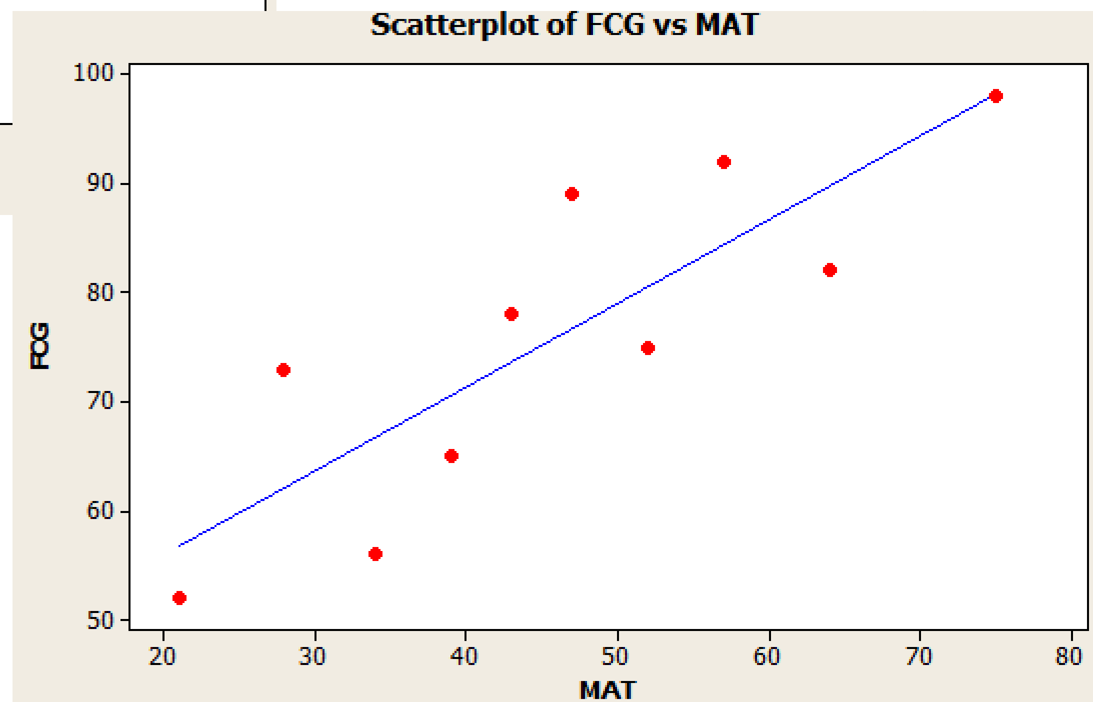


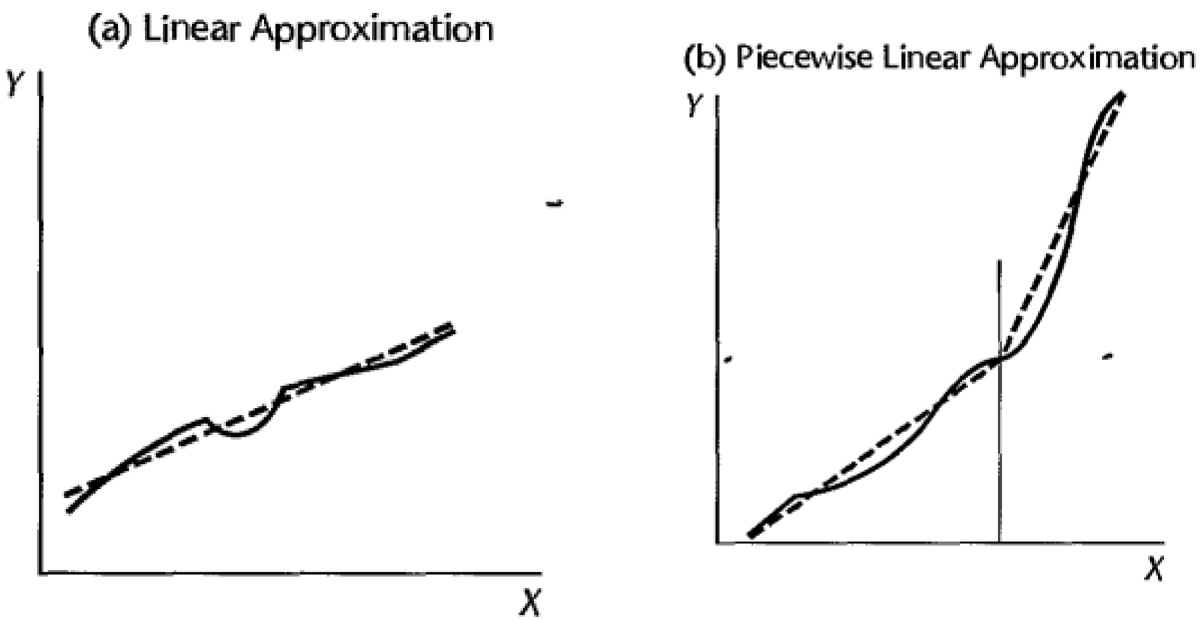
No 2

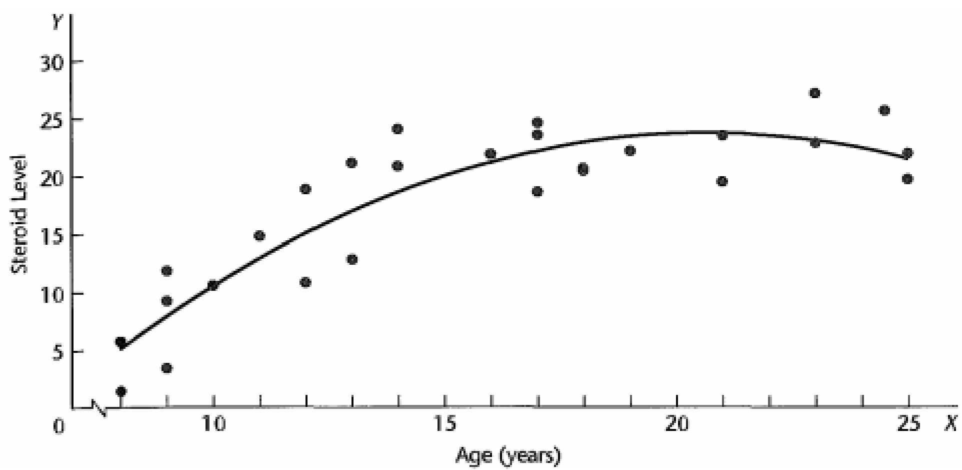
Tabel 1

Mhs	1	2	3	4	5	6	7	8	9	10
MAT	39	43	21	64	57	47	28	75	34	52
FCG	65	78	52	82	92	89	73	98	56	75






(a) Linear Approximation
Y
X
(b) Piecewise Linear Approximation
Y
X


Y
30
25
20
15
10
5
0
Steroid Level
10
15
20
25
X
Age (years)

An Introduction

Regresi linier sering digunakan untuk melihat nilai prediksi atau perkiraan yang akan datang

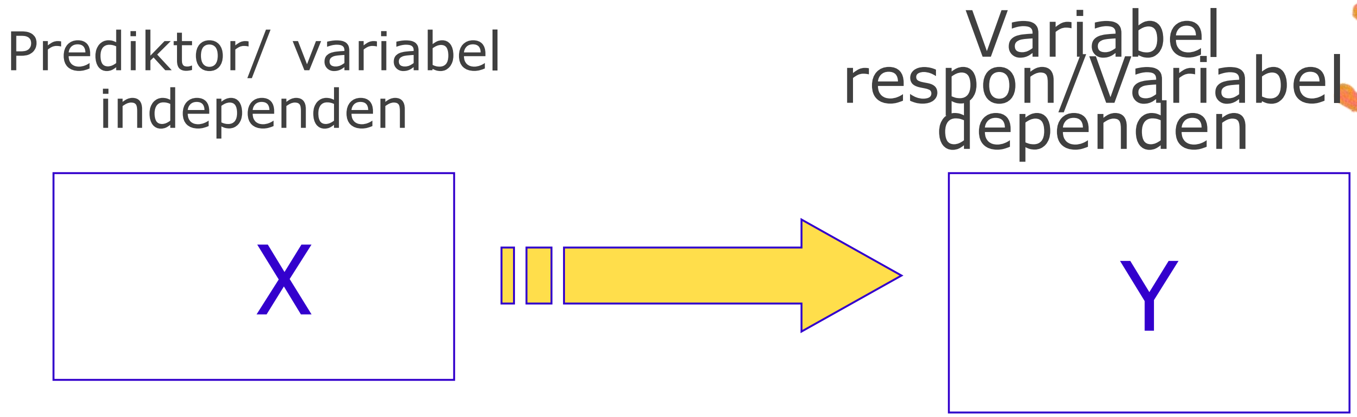
Apabila X dan Y mempunyai hubungan, maka nilai X yang sudah diketahui dapat digunakan memperkirakan Y

→**Variable Y yang nilainya akan diramalkan disebut variabel tidak bebas / variabel respon (*dependent variable*)**

→ **Variable X yang nilainya digunakan untuk meramalkan nilai Y disebut variable bebas/ peramal/ menerangkan (*independent / explanatory variable*)**

Beberapa contoh penelitian mengg. anreg..

- **Pengaruh Efikasi Diri Terhadap Stres Mahasiswa yang Sedang Menyusun Seminar Makalah di Pendidikan Matematika UNS**
- **ANALISIS FAKTOR – FAKTOR YANG MEMPENGARUHI KEBERHASILAN MAHASISWA P. MATEMATIKA UNS**
- **Pengaruh Gaya Kepemimpinan dan Kreativitas Dosen di Kelas terhadap Prestasi Belajar Mahasiswa**
- **PENGARUH KEAKTIFAN DALAM KEGIATAN UKM TERHADAP SOFTSKILL DAN PRESTASI MAHASISWA**



Dapatkah variabel X memprediksi Y ?
Adakah korelasi antara X dan Y?

Analisis regresi digunakan untuk mengetahui bagaimana variabel dependen atau kriterium dapat diprediksikan melalui variabel independen atau prediktor secara individu atau parsial maupun secara bersama-sama atau simultan.

Misal..Ilustrasi hubungan positif antara x dan y

X		Y
Pupuk Berat Badan Keaktifan Kepemimpinan Kinerja	→	Produksi Tekanan darah Prestasi softskill Produktifitas

Ilustrasi hubungan negatif

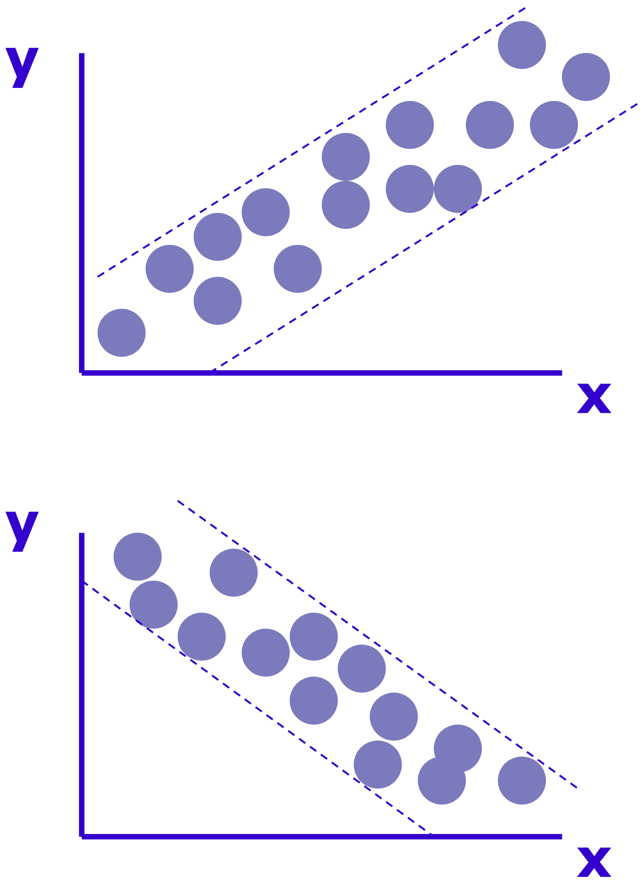
X
Jumlah aseptor
Harga suatu barang

→

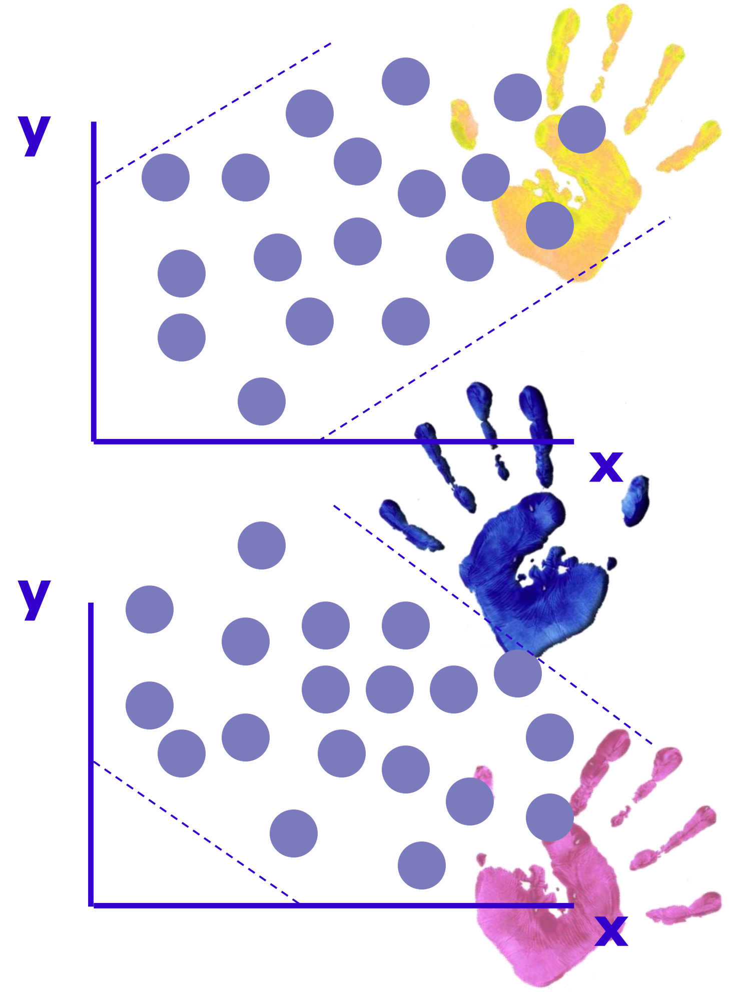
Y
Jumlah kelahiran
Permintaan barang

Contoh plot korelasi

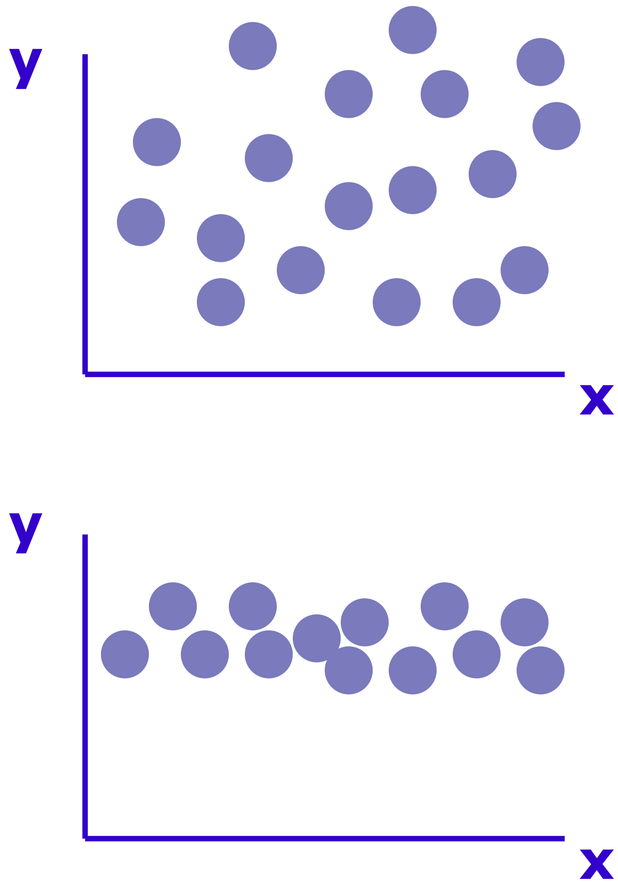
Hubungan kuat



Hubungan lemah



Tidak ada hubungan



I. Regresi linier jika hubungan antara variabel bebas terhadap variabel tak bebas berbentuk linier

- ❑Regresi linier sederhana → $\hat{Y} = a + bX$
- ❑ Regresi linier berganda → $\hat{Y} = a + b_1 X_1 + b_2 X_2 + b_3 X_3$
- ❑Regresi Logistik (Netter :555)
- ❑Regresi Poisson

II. Regresi tak linier jika hubungan antara variabel bebas terhadap variabel tak berbentuk linier

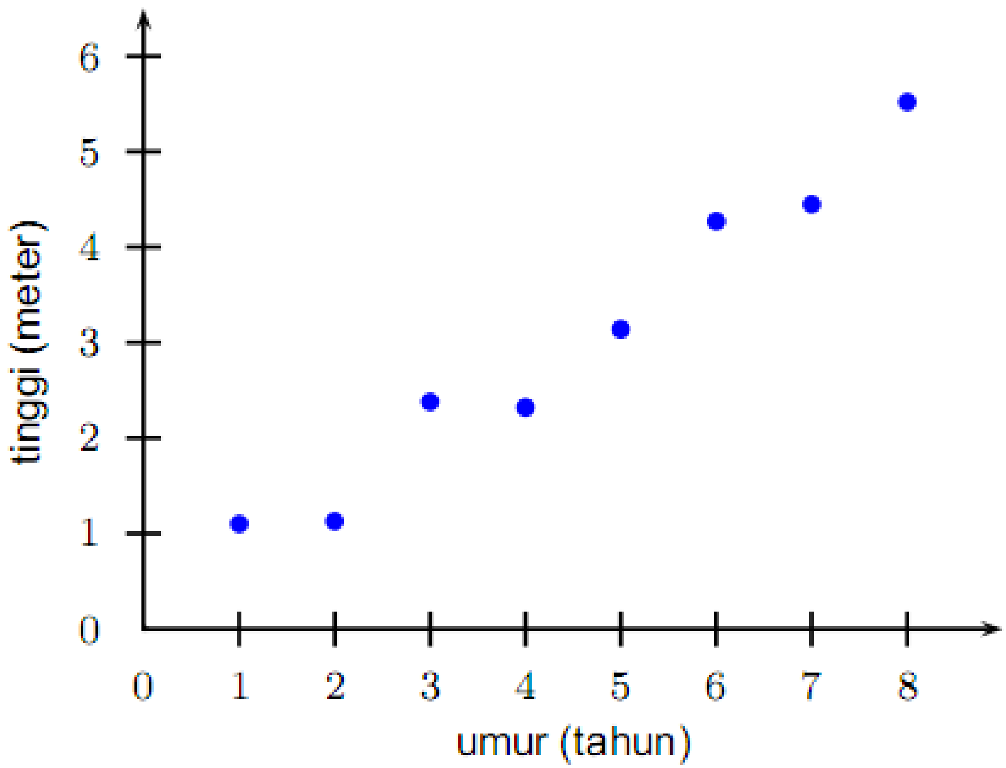
- ❑ Regresi Polinomial → $\hat{Y} = a + bX + cX^2$

$$\hat{Y} = a + bX + cX^2 + dX^3$$

- ❑Neural Network Model (netter : 547)

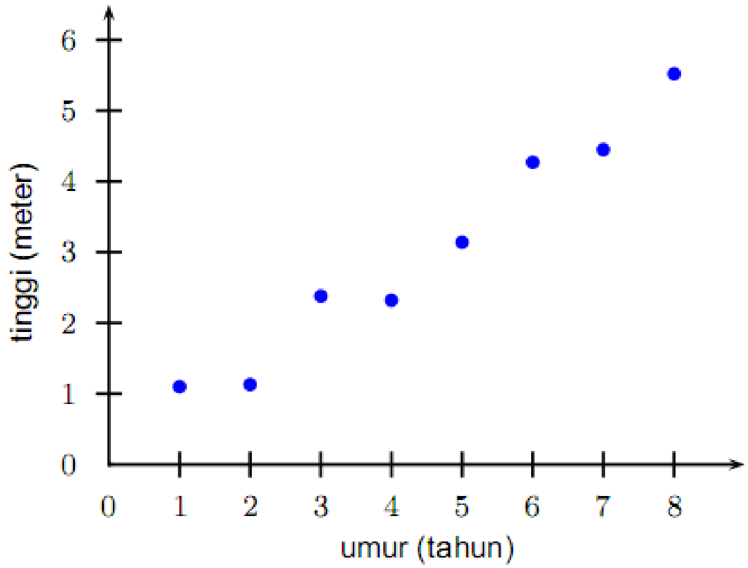
Regresi Linier Sederhana

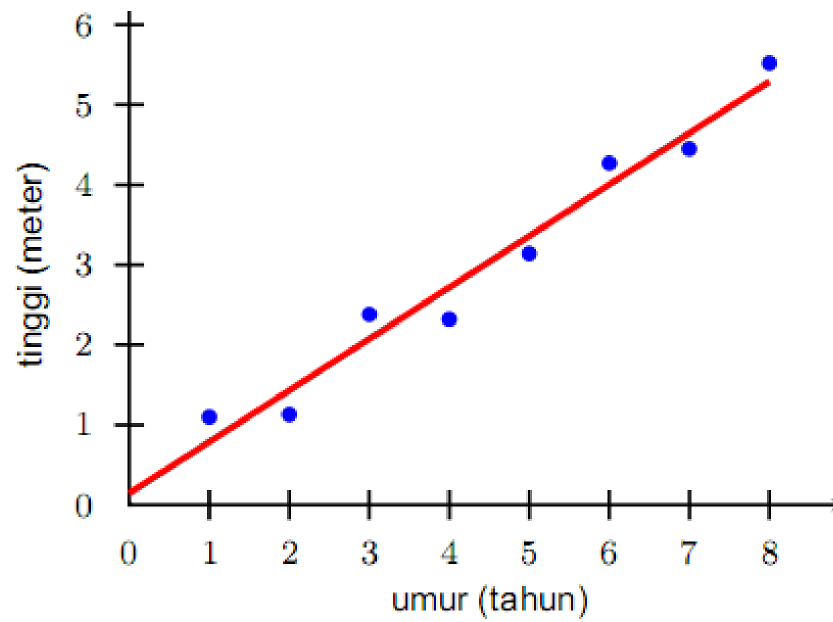
Akan dicari garis linear $\hat{y} = a + bx$ yang paling "mewakili" hubungan antara x (umur) dan y (tinggi)



Dipunyai data umur dan tinggi dari sampel 8 buah pohon jenis tertentu sbb.:

umur (tahun):	1	2	3	4	5	6	7	8
tinggi (meter):	1,10	1,13	2,38	2,32	3,14	4,27	4,45	5,52





Memilih persamaan Terbaik ..?

- Metode Seleksi Maju
- Metode Penyisihan
- Metode Bertahap
- Metode R-square maksimum (MAXR)
- Metode PRESS

Sembiring, 1995

X_i : variabel independen ke-i

Y_i : variabel dependen ke-i maka bentuk model regresi sederhana adalah :

$$Y_i = \alpha + \beta X_i + \varepsilon_i, \quad i = 1,2,\ldots,n$$

dengan

$\hat{\alpha}, \hat{\beta}$ atau a, b parameter yang tidak diketahui

ε_i sesatan random dgn asumsi

NID $(0, \sigma^2)$

$$E[\varepsilon_i] = 0, Var(\varepsilon_i) = \sigma^2$$

$$\Downarrow$$

ε_i dan ε_j tidak saling berkorelasi sehingga kovariansinya sama dengan 0,

$$\sigma\{\varepsilon_i, \varepsilon_j\} = 0, \forall i, j, i \neq j, i = 1,2,\ldots,n$$

$$Y_i = \alpha + \beta X_i + \varepsilon_i, \quad i = 1,2,\ldots,n$$

$$E[Y_i] = E[\alpha + \beta X_i + \varepsilon_i]$$

$$= \hat{\alpha} + \hat{\beta} X + E[\varepsilon_i]$$

So...

$$\hat{Y}_i = a + bX$$

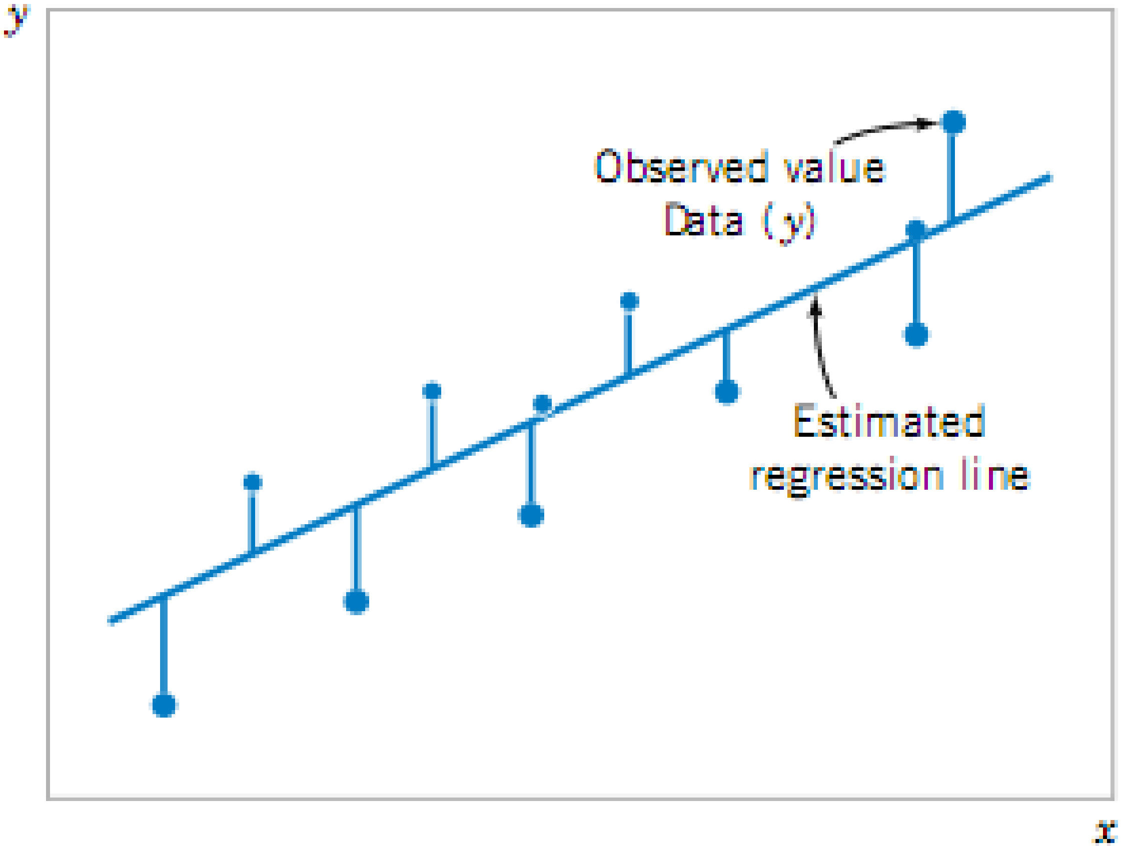
$$Y_i = \alpha + \beta X_i + \varepsilon_i, \quad i = 1,2,\ldots,n$$

$$V(Y_i) = V(\alpha + \beta X_i + \varepsilon_i)$$

$$= V(\alpha + \beta X_i) + V(\varepsilon_i)$$

So...

$$V(Y_i) = 0 + \sigma^2$$



Dari garis regresi sampel diperoleh :

$$e_i = Y_i - (\hat{\alpha} + \hat{\beta} X_i)$$

Dan $$D = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (Y_i - (a + bX_i))^2$$

Turunkan D terhadap a dan b !!!!

$$\frac{\partial D}{\partial a} = -2\sum_{i=1}^{n}\left(Y_i - a - bX_i\right) = 0$$

$$\frac{\partial D}{\partial b} = -2\sum_{i=1}^{n}\left(Y_i - a - bX_i\right)X_i = 0$$

$$\sum_{i=1}^{n} X_i Y_i - a\sum_{i=1}^{n} X_i - b\sum_{i=1}^{n} X_i^2 = 0$$

$$\sum_{i=1}^{n} Yi - an - b\sum_{i=1}^{n} X_i = 0$$

$$\sum_{i=1}^{n} Yi - b\sum_{i=1}^{n} X_i = an$$

$$a = \sum_{i=1}^{n}\frac{Yi}{n} - b\sum_{i=1}^{n}\frac{X_i}{n}$$

$$= \overline{Y} - b\overline{X}$$

$$a = \frac{\sum y_i \sum x_i^2 - \sum x_i \sum x_i y_i}{n\sum x_i^2 - (\sum x_i)^2}$$

$$b = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{n\sum x_i^2 - (\sum x_i)^2}$$

ATAU

y	x	xy	x^2	y^2
: :	: :	: :	: :	: :
Σy	Σx	Σxy	Σx^2	Σy^2

$$b = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - (\sum x)^2}$$

$$a = \bar{y} - b\bar{x}$$

$$\bar{y} = \frac{\sum y}{n} \qquad \bar{x} = \frac{\sum x}{n}$$

Latihan

Carilah persamaan regresi Y pada X dari data Tabel :

Mat (X)	Fis (Y)	XY	X2	Y2
60	80	4800	3600	6400
45	69	3105	2025	4761
50	71	3550	2500	5041
60	85	5100	3600	7225
50	80	4000	2500	6400
65	82	5330	4225	6724
60	89	5340	3600	7921
65	93	6045	4225	8649
50	76	3800	2500	5776
65	86	5590	4225	7396
45	71	3195	2025	5041
50	69	3450	2500	4761
665	951	53305	37525	76095

$$b = \frac{\sum xy - \frac{(\sum x)(\sum y)}{n}}{\sum x^2 - \frac{(\sum x)^2}{n}}$$

$$= \frac{53305 - \frac{(665)(951)}{12}}{37525 - \frac{(665)^2}{12}}$$

$$= 0.8972$$

$$a = \bar{y} - b_1 \bar{x} = 29.53$$

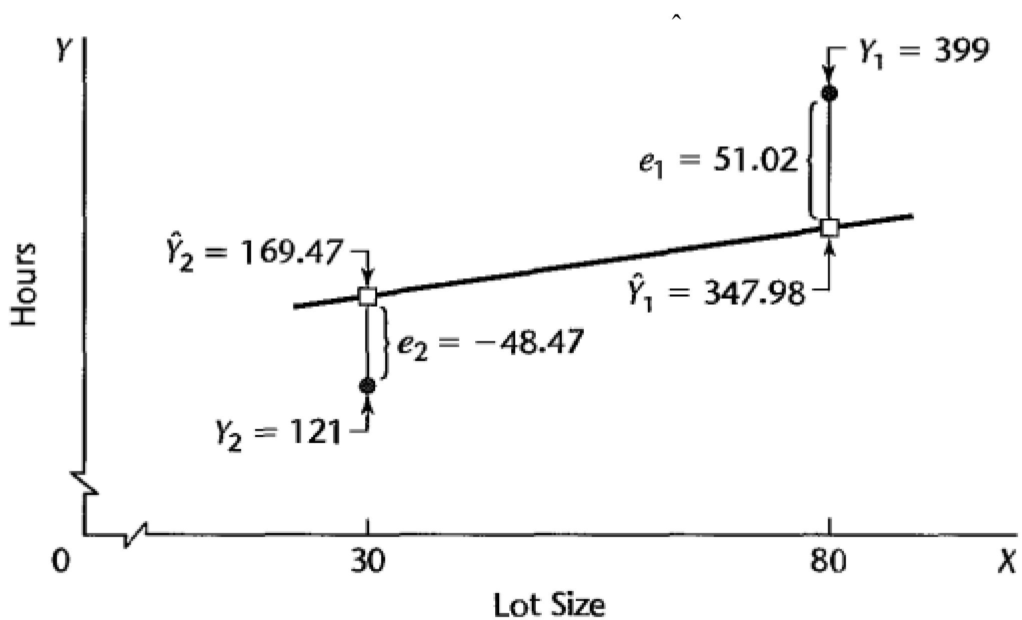
jadi diperoleh persamaan regresi :

$$\hat{Y}_i = 29.5294 + 0.8972\,X_i$$

Residual

- Merupakan nilai selisih antara Yi dengan prediksi Yi, dinotasikan dengan ei,

Contoh ilustrasi



Sifat residual

1. $\sum_{i=1}^{n} e_i = 0$

2. $\sum_{i=1}^{n} e_i^2$ minimum

Type equation here.

3. $\sum_{i=1}^{n} Y_i = \sum_{i=1}^{n} \hat{Y}_i$

4. $\sum_{i=1}^{n} X_i e_i = 0$

5. $\sum_{i=1}^{n} \hat{Y}_i e_i = 0$

6. garis regresi selalu melewati titik $\left(\overline{X}, \overline{Y}\right)$

Latihan

Airfreight breakage. A substance used in biological and medical research is shipped by airfreight to users in cartons of 1,000 ampules. The data below, involving 10 shipments, were collected on the number of times the carton was transferred from one aircraft to another over the shipment route (X) and the number of ampules found to be broken upon arrival (Y). Assume that first-order regression model (1.1) is appropriate.

i:	1	2	3	4	5	6	7	8	9	10
X_i:	1	0	2	0	3	1	0	1	2	0
Y_i:	16	9	17	12	22	13	8	15	19	11

a. Obtain the estimated regression function. Plot the estimated regression function and the data. Does a linear regression function appear to give a good fit here?

b. Obtain a point estimate of the expected number of broken ampules when $X = 1$ transfer is made.

c. Estimate the increase in the expected number of ampules broken when there are 2 transfers as compared to 1 transfer.

d. Verify that your fitted regression line goes through the point $(\bar{X}, \bar{Y})$.