



BAB 3

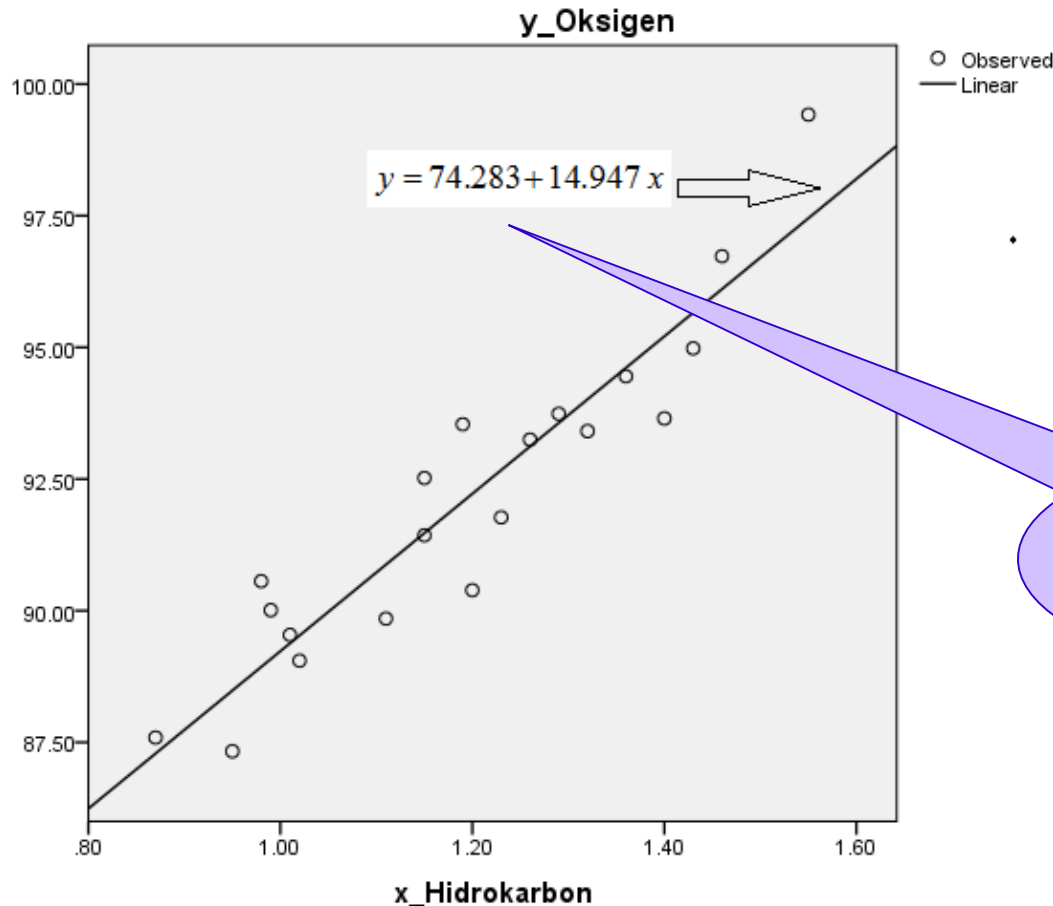
ANALISIS REGRESI



No 1.

Contoh aplikasi Anreg di Bidang Kimia Industri

Kemurnian oksigen (y) yang dihasilkan dari proses distilasi sedangkan x merupakan persentase hidrokarbon yang terkandung dalam kondenser distilasi. Ingin diketahui ada tidaknya pengaruh x terhadap y.



Pendekatan model
regresi linier

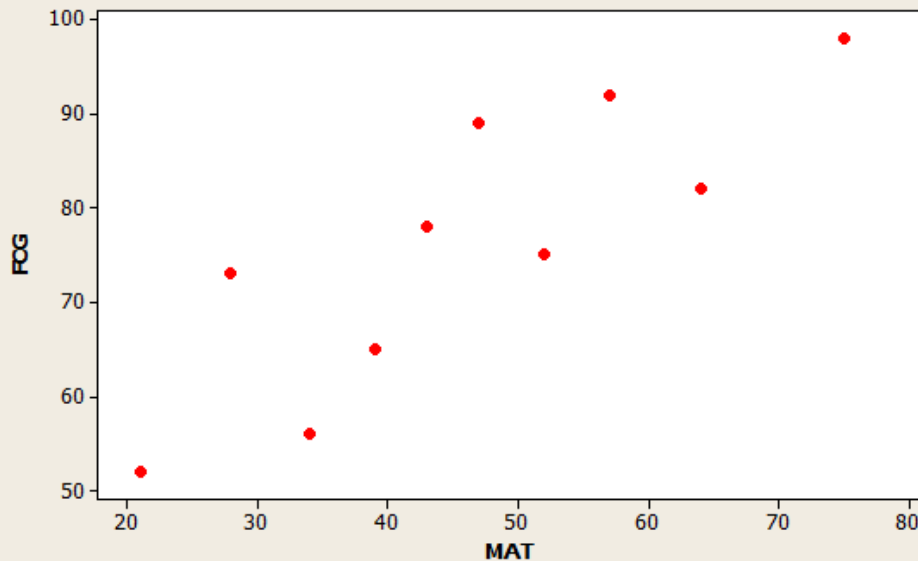
No 2. Penelitian anreg untuk mengetahui ada tidaknya pengaruh emosional (FCG) terhadap prestasi matematika

Tabel 1

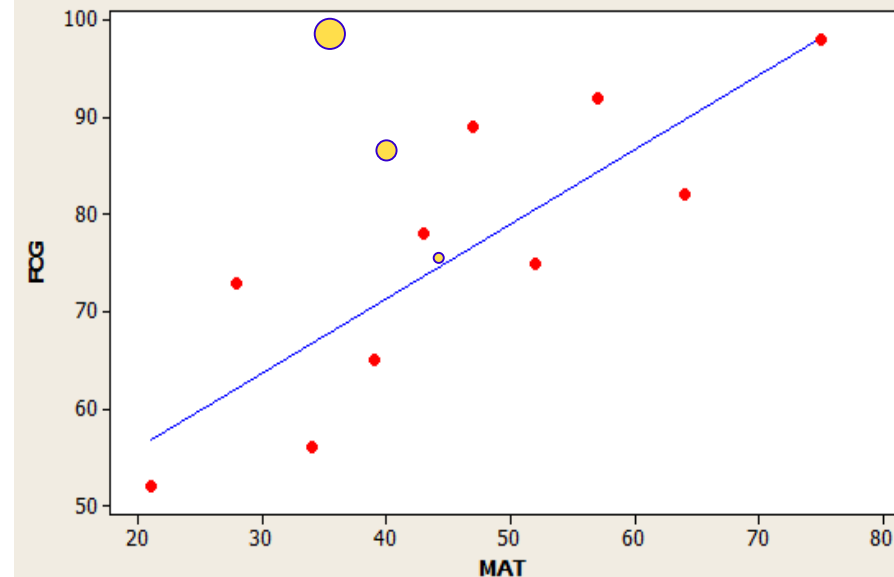
<u>Mhs</u>	1	2	3	4	5	6	7	8	9	10
<u>MAT</u>	39	43	21	64	57	47	28	75	34	52
<u>FCG</u>	65	78	52	82	92	89	73	98	56	75

Garis regresi

Scatterplot of FCG vs MAT

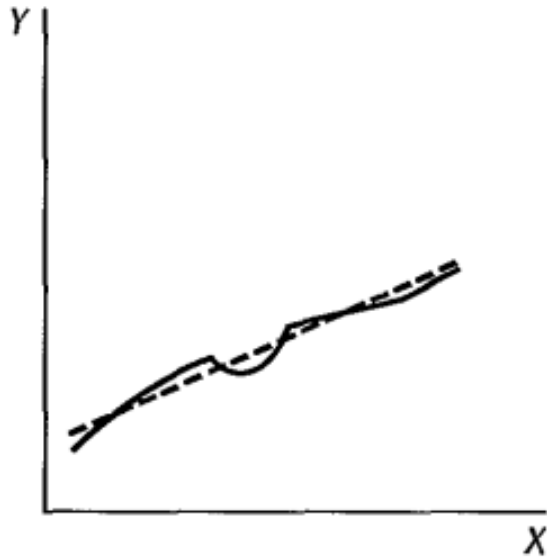


Scatterplot of FCG vs MAT

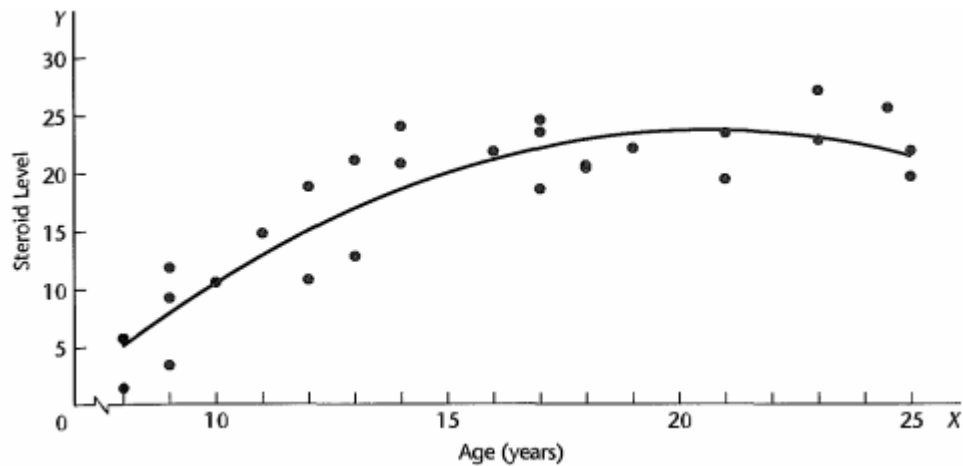
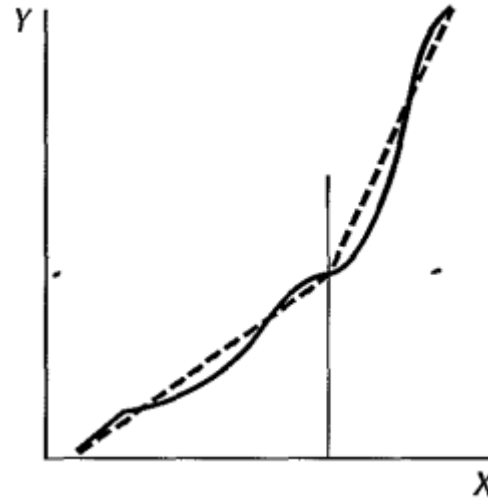


PENDEKATAN model ...

(a) Linear Approximation



(b) Piecewise Linear Approximation



PENGANTAR ANREG..

- Variable Y yang nilainya akan diramalkan disebut variabel tidak bebas / variabel respon (*dependent variable*)
- Variable X yang nilainya digunakan untuk meramalkan nilai Y disebut variabel bebas/ peramal/ menerangkan (*independent / explanatory variable*)
- Apabila X dan Y mempunyai hubungan, maka nilai X yang sudah diketahui dapat digunakan memperkirakan Y

ANREG

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

X



Y



Sebutkan contoh simulasi peneltiang menggunakan Anreg di Pendidikan Matematika ...

1. Y :
2. Y :
3. Y :



Misal..Ilustrasi hubungan positif antara x dan y

X
Pupuk
Berat Badan
Keaktifan
Kepemimpinan
Kinerja



Y
Produksi
Tekanan darah
Prestasi
softskill
Produktifitas



Misal..Ilustrasi hubungan negatif antara x dan y

X
Jumlah aseptor
Harga suatu barang

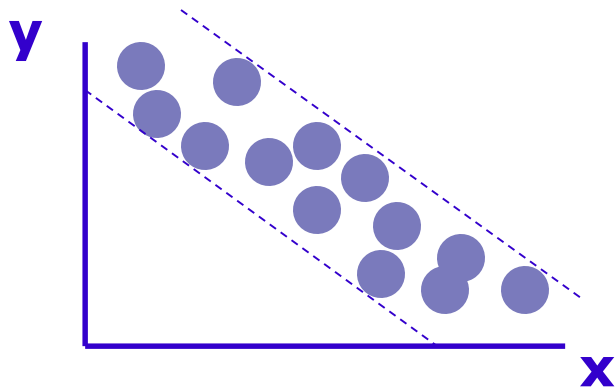
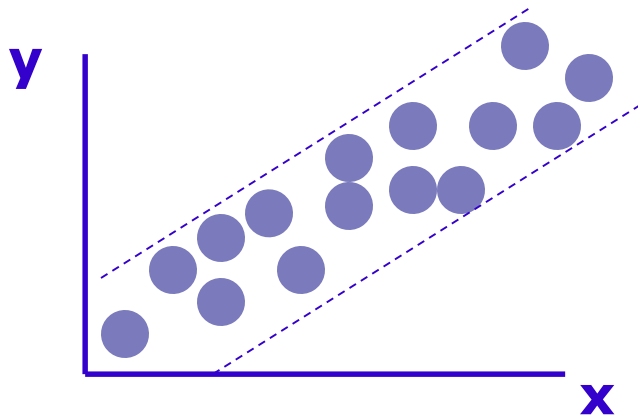


Y
Jumlah kelahiran
Permintaan barang

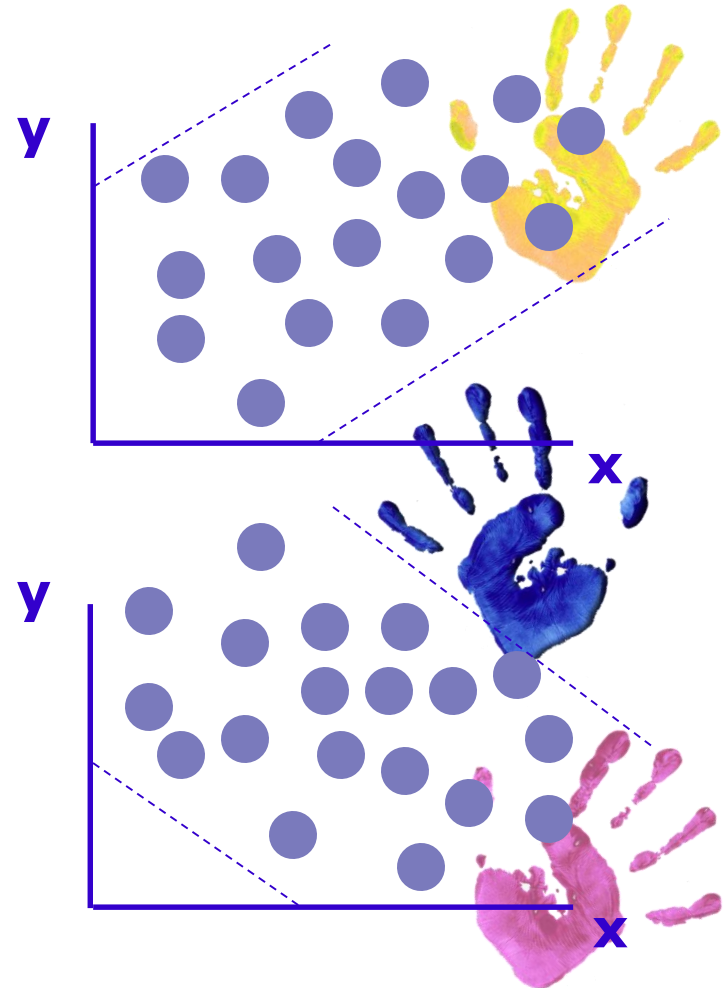


Contoh plot korelasi

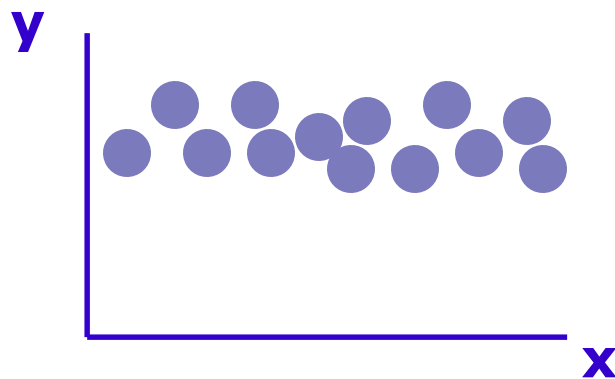
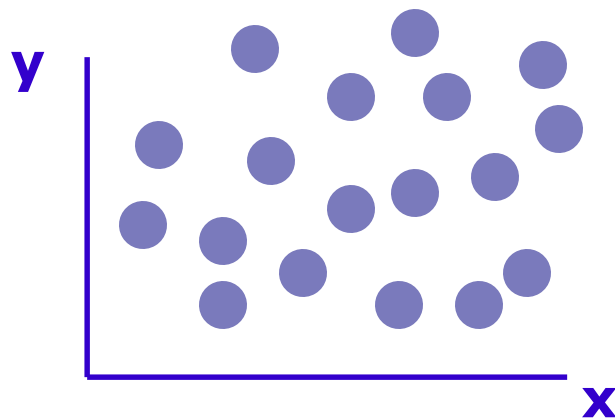
Hubungan kuat



Hubungan lemah



Tidak ada hubungan



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

- ❑ Regresi linier sederhana $\rightarrow \hat{Y} = a + bX$
- ❑ Regresi linier berganda $\rightarrow \hat{Y} = a + b_1X_1 + b_2X_2 + b_3X_3$
- ❑ Regresi Logistik (Netter : 555)
- ❑ Regresi Poisson

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

- ❑ Regresi Polinomial $\rightarrow \hat{Y} = a + bX + cX^2$
 $\hat{Y} = a + bX + cX^2 + dX^3$
- ❑ **Neural Network Model** (netter : 547)



Regresi Linier Sederhana

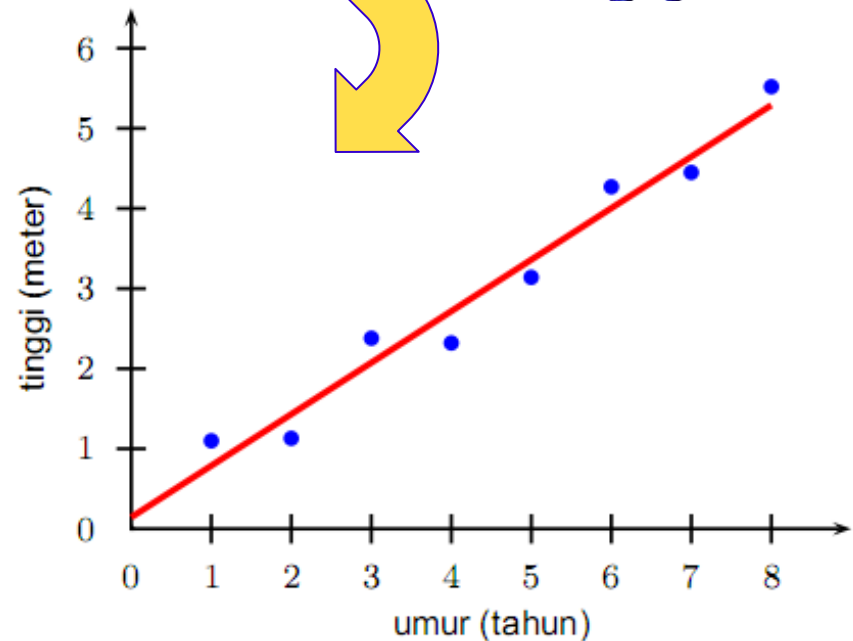
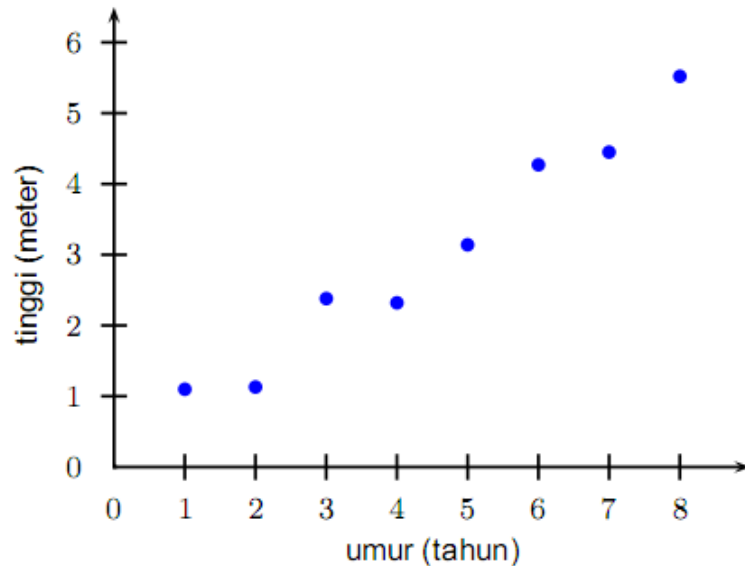


Akan dicari garis linear $\hat{y} = a + bx$ yang paling "mewakili" hubungan antara x (umur) dan y (tinggi)

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



X_i : variabel independen ke- i

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

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

dengan

α, β parameter yang tidak diketahui

ε_i sesatan random dgn asumsi

NID $(0, \sigma^2)$

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

$\Downarrow$

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

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

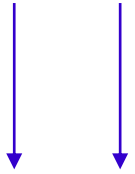


$$Y_i = \alpha + \beta X_i + \varepsilon_i, \quad i = 1, 2, \dots, n$$

$$\begin{aligned} E[Y_i] &= E[\alpha + \beta X_i + \varepsilon_i] \\ &= \hat{\alpha} + \hat{\beta}X + E[\varepsilon_i] \end{aligned}$$

So...

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



Koefisien regresi (di sampel)
Yang akan diestimasi

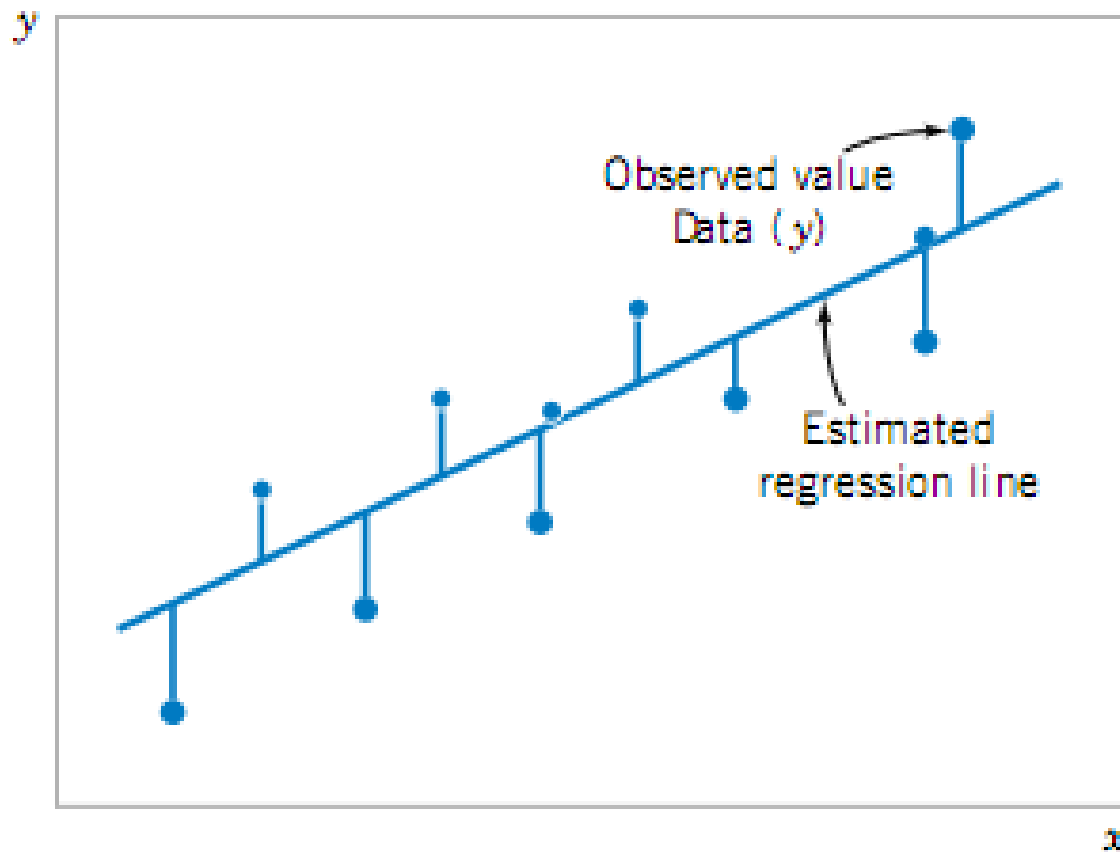
$$Y_i = \alpha + \beta X_i + \varepsilon_i, \quad i = 1, 2, \dots, n$$

$$\begin{aligned} V(Y_i) &= V(\alpha + \beta X_i + \varepsilon_i) \\ &= V(\alpha + \beta X_i) + V(\varepsilon_i) \end{aligned}$$

So...

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





Dari garis regresi sampel diperoleh :

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

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

Dengan MKT, untuk mencari nilai estimasi a dan b maka Turunkan D terhadap a dan b !!!!

Turunan terhadap a

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

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

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

$$a = \sum_{i=1}^n \frac{Y_i}{n} - b \sum_{i=1}^n \frac{X_i}{n}$$
$$= \bar{Y} - b\bar{X}$$

dengan

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

Turunan terhadap b

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

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

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

Atau

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

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



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:

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

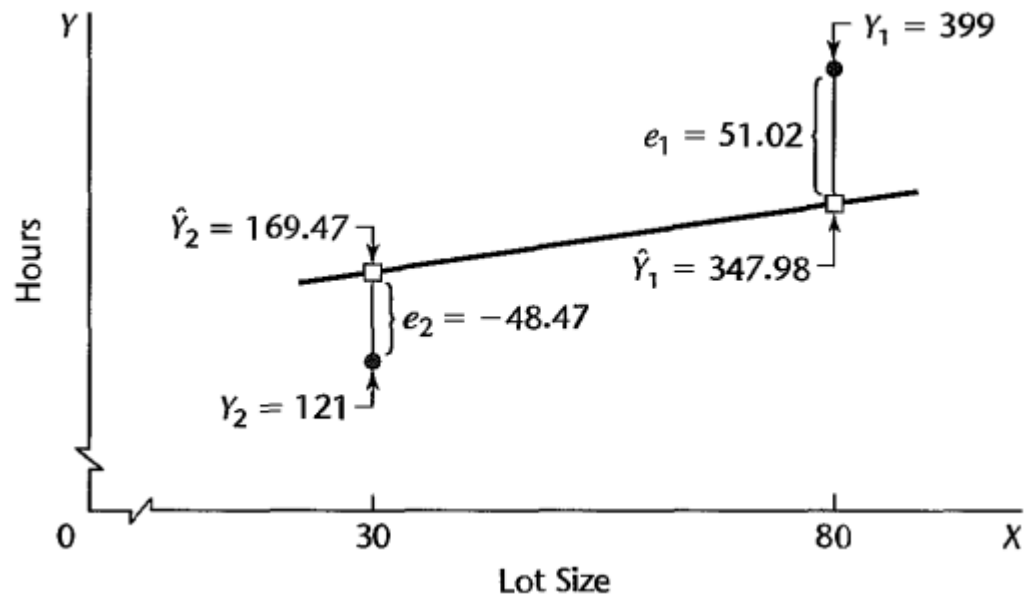
Coefficients ^a								
Model	Unstandardized Coefficients			Standardized Coefficients	t	Sig.	Collinearity Statistics	
	B	Std. Error		Beta			Tolerance	VIF
1 (Constant)	29.529	9.311			3.171	.010		
matematik	.897	.167		.862	5.389	.000	1.000	1.000

a. Dependent Variable: fisika

Residual

Merupakan nilai selisih antara Y_i dengan prediksi $\hat{Y}_i$, dinotasikan dengan e_i ,

- Contoh



- Sifat

1. $\sum_{i=1}^n e_i = 0$
2. $\sum_{i=1}^n e_i^2$ minimum
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 $(\bar{X}, \bar{Y})$



Latihan

Tabel 4.3 merupakan data yang berisi kemurnian oksigen (y) yang dihasilkan dari proses distilasi sedangkan x merupakan persentase hidrokarbon yang terkandung dalam kondenser distilasi. Peneliti ingin mengetahui ada tidaknya hubungan antara hidrokarbon terhadap kemurnian oksigen yang dihasilkan dalam proses distilasi (misalkan digunakan tingkat signifikansi 5%)



1	2	3	1	2	3
1	0.99	90.01	11	1.19	93.54
2	1.02	89.05	12	1.15	92.52
3	1.15	91.43	13	0.98	90.56
4	1.29	93.74	14	1.01	89.54
5	1.46	96.73	15	1.11	89.85
6	1.36	94.45	16	1.2	90.39
7	0.87	87.59	17	1.26	93.25
8	1.23	91.77	18	1.32	93.41
9	1.55	99.42	19	1.43	94.98
10	1.4	93.65	20	0.95	87.33

keterangan:

Kolom 1 : No observasi

Kolom 2 : Hidrokarbon_x (%)

Kolom 3 : Oksigen_y