

ANALISIS REGRESI DAN KORELASI



TEKNIK INFORMATIKA
APRIANI PUTI PURFINI,S.KOM.,M.T.

Analisis Korelasi

Hubungan Antar Peubah

- Dari setiap objek / individu / tempat / dll dapat diukur / dicatat / diamati lebih dari satu buah peubah
- Nilai dari suatu peubah bersifat :
 - saling bebas dengan peubah lain
 - saling terkait dengan peubah lain

Analisis Korelasi

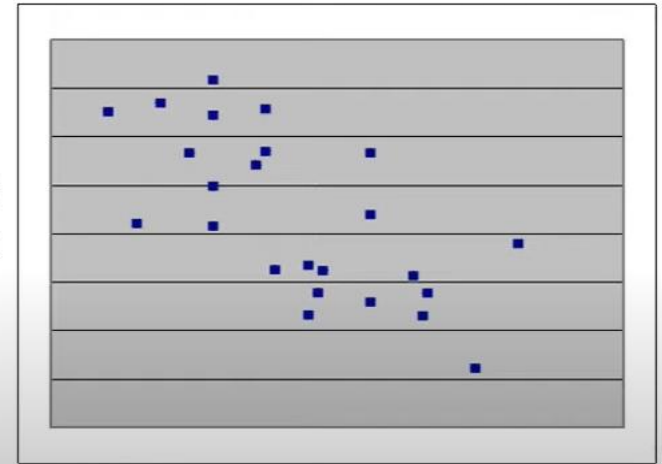
Hubungan Antar Peubah

berat
badan



tinggi badan

suhu rata-
rata



ketinggian tempat

Analisis Korelasi

Koefesien Korelasi

- Sebuah ukuran yang dapat mencirikan keeratan hubungan antar dua peubah
- Koefisien Korelasi (ρ ; baca : rho)
 - Nilainya : $-1 \leq \rho \leq 1$
 - Tanda menunjukkan arah hubungan
 - Besar / magnitude menunjukkan kekuatan hubungan
 - Koefesien korelasi data contoh dinotasikan r

Analisis Korelasi

Koefesien Korelasi (Peubah)

Jika ada dua peubah X dan Y, korelasi antara keduanya adalah

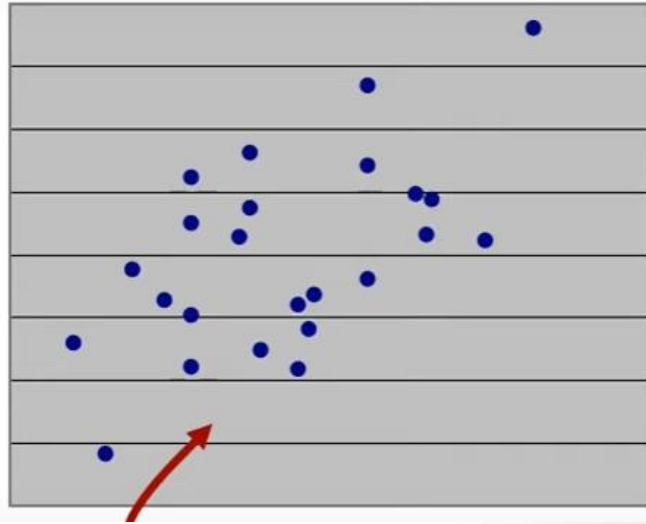
$$r_{xy} = \frac{S_{xy}}{S_x S_y}$$

$$S_{xy} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{n - 1}$$

$$S_x = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n - 1}} \quad \text{dan} \quad S_y = \sqrt{\frac{\sum (y_i - \bar{y})^2}{n - 1}}$$

Analisis Korelasi

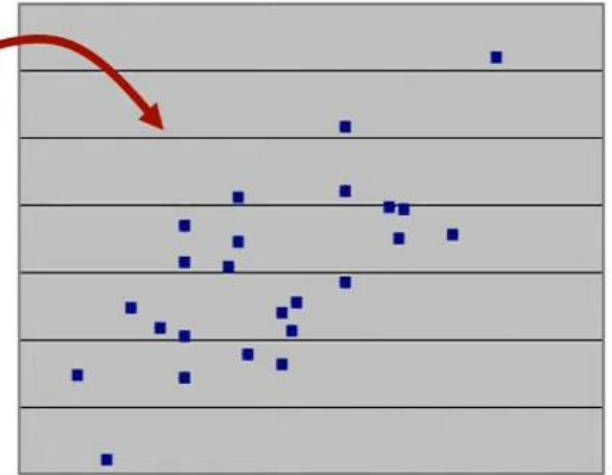
Koefesien Korelasi (+)



$r = 0.58$



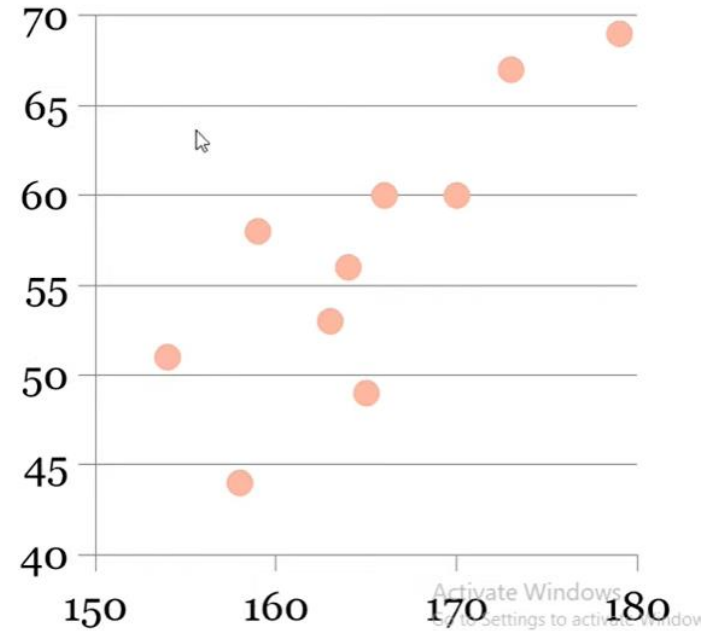
$r = 0.70$



Analisis Korelasi

Ilustrasi

Tinggi Badan	Berat Badan
165	49
159	58
166	60
173	67
179	69
164	56
163	53
154	51
170	60
158	44



Analisis Korelasi

Contoh

X	y	x-xbar	y-ybar	(x-xbar) ²	(y-ybar) ²	(x-xbar)(y-ybar)
165	49	-0.1	-7.7	0.01	59.29	0.77
159	58	-6.1	1.3	37.21	1.69	-7.93
166	60	0.9	3.3	0.81	10.89	2.97
173	67	7.9	10.3	62.41	106.09	81.37
179	69	13.9	12.3	193.21	151.29	170.97
164	56	-1.1	-0.7	1.21	0.49	0.77
163	53	-2.1	-3.7	4.41	13.69	7.77
154	51	-11.1	-5.7	123.21	32.49	63.27
170	60	4.9	3.3	24.01	10.89	16.17
158	44	-7.1	-12.7	50.41	161.29	90.17
rata-rata						
165.1	56.7	jumlah		496.9	548.1	426.3

$$r_{xy} = \frac{S_{xy}}{S_x S_y}$$

$$S_{xy} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{n - 1} = \frac{426.3}{10 - 1} = 47.25$$

$$S_x = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n - 1}} = \sqrt{\frac{496.9}{10 - 1}} = 7.4304$$

$$S_y = \sqrt{\frac{\sum (y_i - \bar{y})^2}{n - 1}} = \sqrt{\frac{548.1}{10 - 1}} = 7.804$$

$$r_{xy} = \frac{S_{xy}}{S_x S_y} = \frac{47.25}{7.4304 \cdot 7.804} = 0.814$$

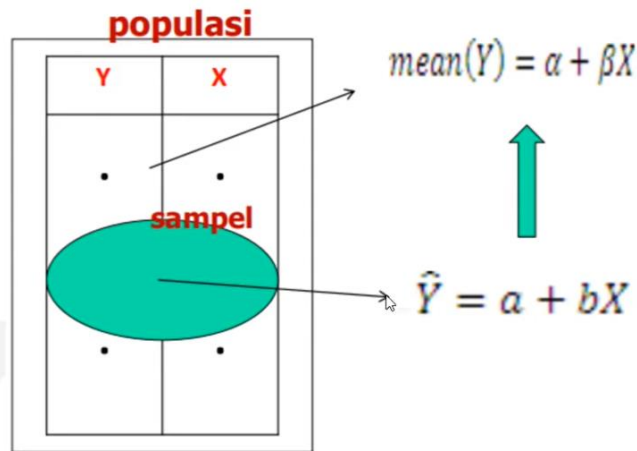
Regresi Linear

Pengantar

Tujuan :

Menyelidiki bentuk atau pola hubungan antara Y dan X.

Mengestimasi / menduga mean atau rata-rata dari Y dari populasi berdasarkan X yang diberikan



Regresi Linear

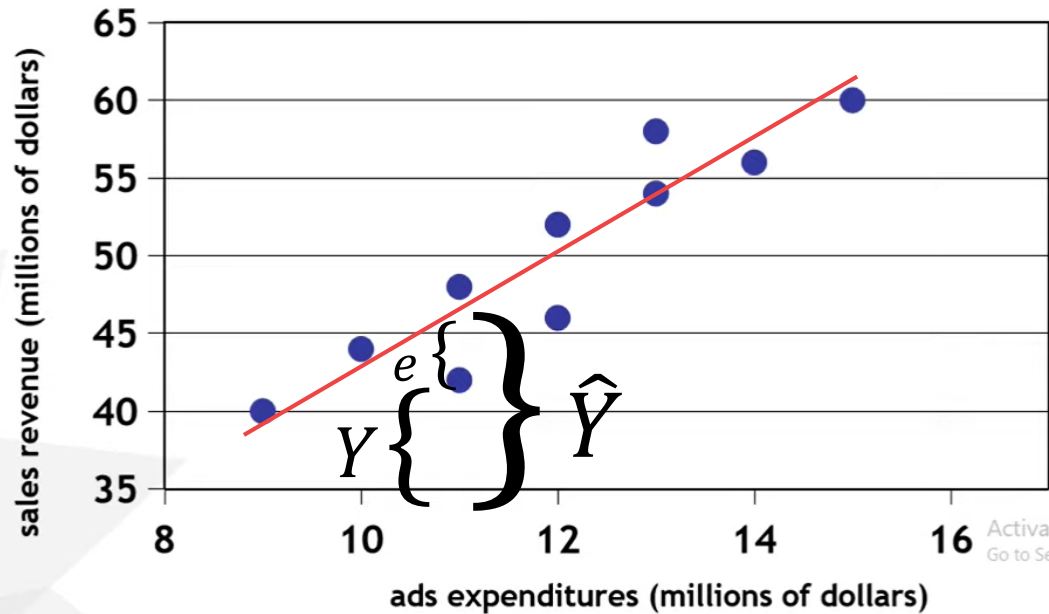
Pengantar

Misalkan ingin melihat hubungan antara pengeluaran untuk iklan (ads expenditures, X) dengan penerimaan melalui penjualan (sales revenue, Y)

Bulan	1	2	3	4	5	6	7	8	9	10
X	10	9	11	12	11	12	13	13	14	15
Y	44	40	42	46	48	52	54	58	56	60

Regresi Linear

Pengantar



Ingin dibuat model

$$Y = a + bX$$

Model memuat error , selisih nilai sebenarnya dengan dugaan berdasar model

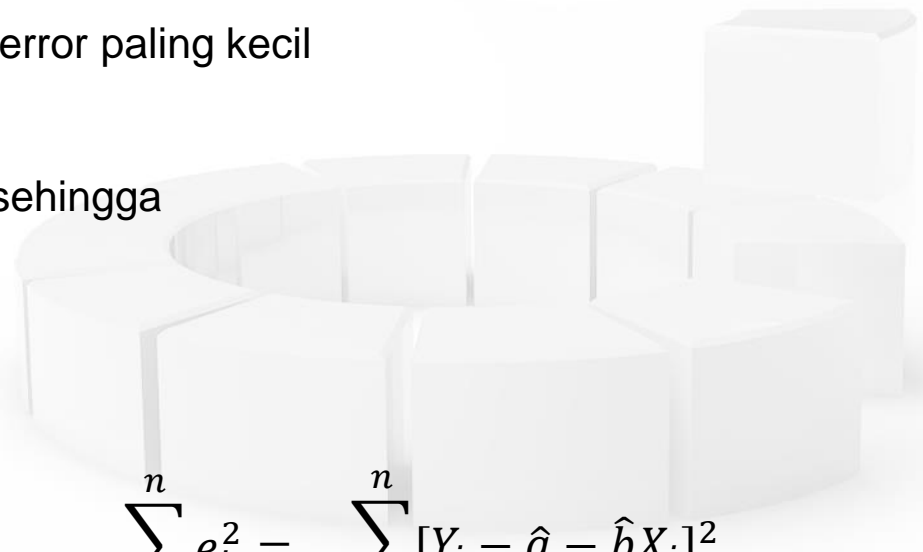
Regresi Linear

Bagaimana mendapatkan a dan b ?

Metode yang digunakan : OLS (ordinary least squares / kuadrat terkecil) , mencari a dan b sehingga jumlah kuadrat error paling kecil

Cari penduga a dan b sehingga

minimum


$$\sum_{i=1}^n e_i^2 = \sum_{i=1}^n [Y_i - \hat{a} - \hat{b}X_i]^2$$

Regresi Linear

Bagaimana mendapatkan a dan b ?

$$\hat{b} = \frac{\sum_{i=1}^n [X_i - \bar{X}][Y_i - \bar{Y}]}{\sum_{i=1}^n [X_i - \bar{X}]^2}$$

$$\hat{a} = \bar{Y} - \hat{b}\bar{X}$$

dimana :

\bar{Y} = rata - rata Y

\bar{X} = rata - rata X

Regresi Linear

Bagaimana mendapatkan a dan b ?

No	X_i	Y_i	$X_i - \bar{X}$	$Y_i - \bar{Y}$	$(X_i - \bar{X})(Y_i - \bar{Y})$	$(X_i - \bar{X})^2$
1	10	44	-2	-6	12	4
2	9	40	-3	-10	30	9
3	11	42	-1	-8	8	1
4	12	46	0	-4	0	0
5	11	48	-1	-2	2	1
6	12	52	0	2	0	0
7	13	54	1	4	4	1
8	13	58	1	8	8	1
9	14	56	2	6	12	4
10	15	60	3	10	30	9
Σ	120	500			106	30

$$\hat{b} = \frac{\sum_{i=1}^n [X_i - \bar{X}][Y_i - \bar{Y}]}{\sum_{i=1}^n [X_i - \bar{X}]^2}$$

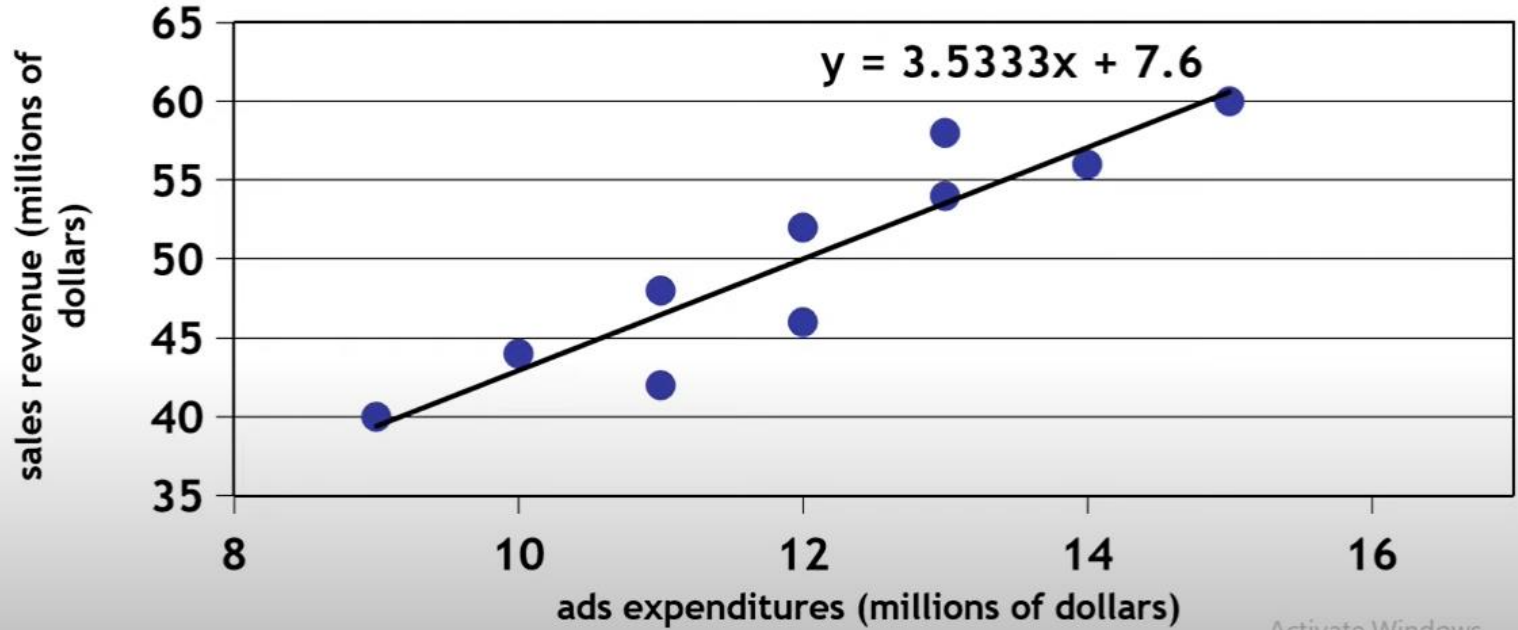
$$\bar{X} = 12 \quad \sum_{i=1}^n [X_i - \bar{X}][Y_i - \bar{Y}] = 106$$

$$\bar{Y} = 50 \quad \sum_{i=1}^n [X_i - \bar{X}]^2 = 30$$

$$\hat{b} = \frac{106}{30} = 3.533$$

$$\hat{a} = \bar{Y} - \hat{b}\bar{X} = 50 - 3.533(12) = 7.60$$

Regresi Linear



Regresi Linear

Intrepretasi a dan b

$$Y = 7.6 + 3.53 X$$

$$Pendapatan = 7.6 + 3.53 \text{ Belanja Iklan}$$

- a = intersep / intercept = besarnya nilai Y ketika X sebesar 0
- b = gradient / slope = besarnya perubahan nilai Y ketika X berubah satu satuan.
Tanda koefesien b menunjukkan arah hubungan X dan Y

Pada kasus ilustrasi

- a = 7.6 → besarnya sales revenue jika tidak ada belanja iklan adalah 7.6 juta dolar
- b = 3.533 → jika belanja iklan dinaikkan 1 juta dolar maka sales revenue naik 3.533 juta dolar