

ANALISA REGRESI & KORELASI SEDERHANA

Analisis Regresi dan Korelasi

- Analisis regresi digunakan untuk mempelajari dan mengukur hubungan statistik yang terjadi antara dua variabel atau lebih variabel.
- Variabel tersebut adalah variabel X (variabel independent / variabel yang mempengaruhi / variabel yang diketahui), dan variabel Y (variabel dependent / variabel yang dipengaruhi/ variabel yang tidak diketahui)
- Analisis korelasi bertujuan untuk mengukur “seberapa kuat” atau “derajat kedekatan”, suatu relasi yang terjadi antar variabel.

Macam hubungan antara 2 variabel

- Pada dasarnya hubungan antar 2 variabel dapat dibedakan atas:
 1. Hubungan searah/positif
 2. Hubungan bersifat kebalikan/negatif
 3. Tidak ada hubungan

Hubungan searah/positif

- Hubungan yang searah diartikan apabila perubahan variabel x (independent) akan mempengaruhi variabel y (dependent) yang searah.

Atau jika variabel x bertambah, maka variabel y bertambah pula, dan sebaliknya.

- Contoh :
 - a. hubungan antara pengeluaran iklan (x) dan jumlah penjualan (y).
 - b. Hubungan antara penghasilan (X) dan pengeluaran konsumsi (Y)

Hubungan bersifat kebalikan/negatif

- Dua variabel dikatakan mempunyai hubungan yang bersifat kebalikan atau negatif, apabila perubahan variabel independent (x) akan mempengaruhi variabel dependent (Y) pada arah yang berlawanan.
- Artinya apabila variabel x bertambah, maka variabel y berkurang atau sebaliknya, jika variabel x berkurang maka variabel y bertambah.

Hubungan bersifat kebalikan/negatif

Contoh :

- a. Hubungan antara usia kendaraan (X) dengan tingkat harga (Y).
- b. Hubungan antara harga barang (x) dengan jumlah yang diminta (Y)

Tidak ada hubungan

- Dua variabel dikatakan tidak punya hubungan apabila perubahan pada variabel independent (x) tidak mempengaruhi perubahan pada variabel dependent (y).
- Contoh :
Hubungan antara konsumsi pangan (x) dengan tingginya gedung (y).

Relasi yang Logis

- Seorang pimpinan perusahaan selalu dihadapkan pada masalah pengambilan keputusan yang berkaitan dengan hal-hal di masa mendatang.
- Untuk pengambilan keputusan yang tepat, maka harus berdasarkan pada data yang diketahui, dihubungkan dengan hal-hal di masa mendatang.
- Seperti kita ketahui, pada semua kejadian, baik kejadian ekonomi maupun lainnya, pasti ada faktor yang menyebabkan terjadinya kejadian-kejadian tersebut (merosotnya hasil penjualan tekstil mungkin disebabkan karena kalah bersaing dengan tekstil impor, merosotnya produksi padi mungkin karena pupuknya berkurang, dan lain sebagainya)

Penggambaran Garis Regresi

Ada 2 cara penggambaran garis regresi :

1. Metode diagram berserak (The scatter diagram)
2. Metode jumlah kuadrat terkecil (The least square's method)

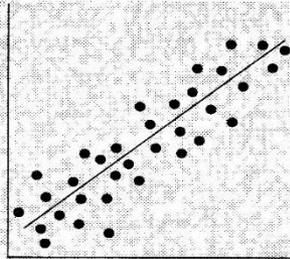
Diagram Pencar

Setelah ditetapkan bahwa terdapat hubungan logis di antara variabel, maka untuk mendukung analisis lebih jauh, barangkali tahap selanjutnya adalah menggunakan **grafik**.

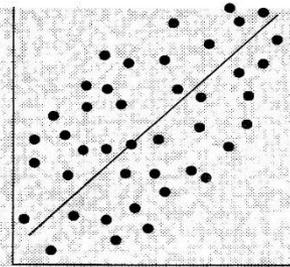
Grafik ini disebut **diagram pencar**, yang menunjukkan titik-titik tertentu. Setiap titik memperlihatkan suatu hasil yang kita nilai sebagai variabel tak bebas maupun bebas

Diagram pencar ini memiliki 2 manfaat, yaitu :

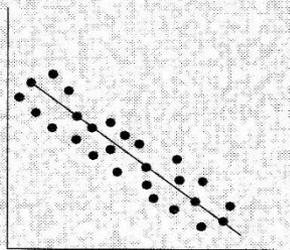
- membantu menunjukkan apakah **terdapat hubungan yang bermanfaat** antara dua variabel,
- dan membantu menetapkan **tipe persamaan** yang menunjukkan hubungan antara kedua variabel tersebut.



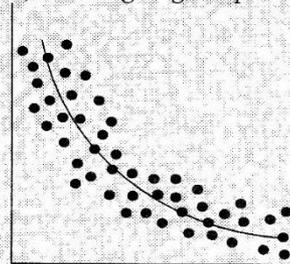
a) Hubungan garis positif



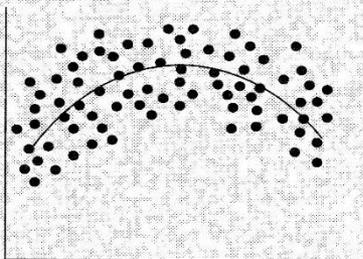
b) Hubungan garis positif tapi lebih tersebar



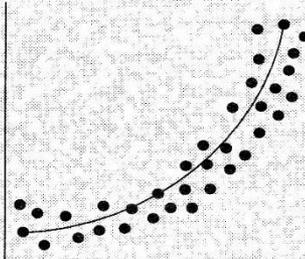
c) Hubungan garis negatif



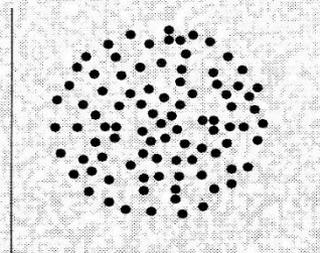
d) Hubungan garis lengkung negatif



e) Hubungan garis lengkung negatif



f) Hubungan garis lengkung positif



g) Tidak ada hubungan

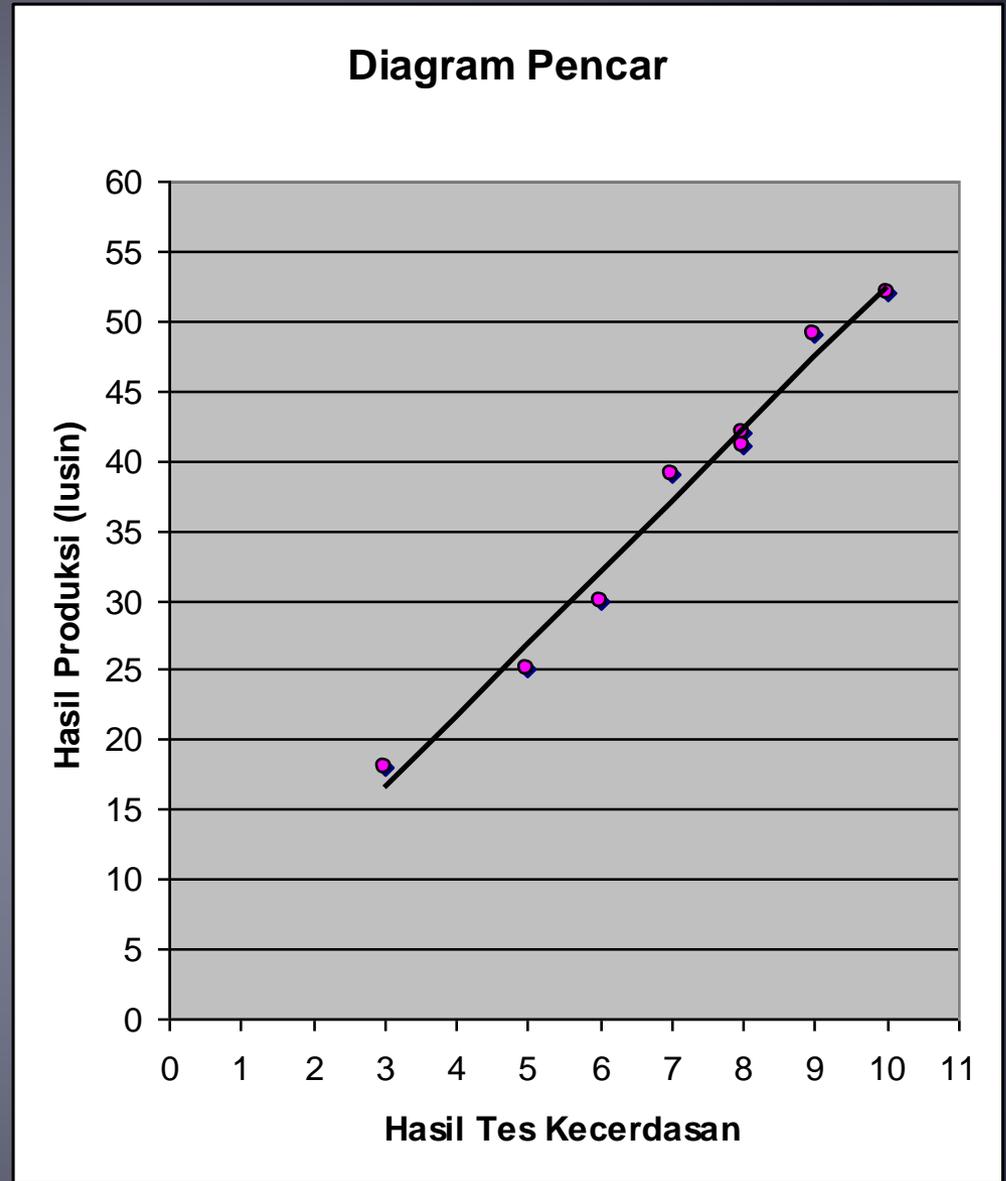
*Dengan jelas, terlihat bahwa data ke 8 karyawan yang kita ambil sebagai sampel adalah kecil, yang memberi kesimpulan tentatif. Tetapi, kita telah memilih sejumlah sampel (contoh) yang kecil untuk meminimalisasi sejumlah keperluan perhitungan pada bagian selanjutnya.

*Billingsley, et. al, *STATISTIC INFERENCE FOR MANAGEMENT AND ECONOMIC*, Third Edition, 1986.

Tabel perhitungan

Karyawan	Hasil Produksi (lusin) (Y)	Skor Tes Kecerdasan (X)
A	30	6
B	49	9
C	18	3
D	42	8
E	39	7
F	25	5
G	41	8
H	52	10

Karyawan	Hasil Produksi (lusin) (Y)	Skor Tes Kecerdasan (X)
A	30	6
B	49	9
C	18	3
D	42	8
E	39	7
F	25	5
G	41	8
H	52	10



Metode jumlah kuadrat terkecil

Regresi merupakan suatu alat ukur yang juga digunakan untuk mengukur ada atau tidaknya korelasi antar variabelnya.

Istilah **regresi** itu sendiri berarti **ramalan** atau **taksiran**.

Persamaan yang digunakan untuk mendapatkan garis regresi pada data diagram pencar disebut persamaan regresi.

Untuk menempatkan garis regresi pada data yang diperoleh maka digunakan metode kuadrat terkecil, sehingga bentuk persamaan regresi adalah sebagai berikut:

$$Y' = a + b X$$

Dimana:

Y' : nilai estimate variabel terikat

a : titik potong garis regresi pd sumbu y (nilai estimate Y' bila $x=0$)

b : gradien garis regresi (perub nilai estimate Y' per satuan perubahan nilai x)

X : nilai variabel bebas

Kesamaan diantara garis regresi dan garis trend tidak dapat berakhir dengan persamaan garis lurus. Garis regresi (seperti garis trend dan nilai tengah aritmatika) memiliki dua sifat matematis berikut :

$$\Sigma(Y - Y') = 0$$

dan $\Sigma(Y - Y')^2 =$ nilai terkecil atau terendah

Dengan perkataan lain, garis regresi akan ditempatkan pada data dalam diagram sedemikian rupa sehingga **penyimpangan** (perbedaan) **positif** titik-titik terhadap titik-titik pencar **di atas** garis akan mengimbangi penyimpangan **negatif** titik-titik pencar yang terletak **di bawah** garis, sehingga hasil penyimpangan keseluruhan titik-titik terhadap garis lurus adalah **nol**.

Untuk tujuan di atas, perhitungan analisis regresi dan analisis korelasi dapat **dipermudah** dengan menggunakan rumus dalam bentuk penyimpangan nilai tengah variabel X dan Y, yaitu penyimpangan dari

$$\bar{X} \text{ dan } \bar{Y}$$

Oleh karena itu, dapat digunakan simbol berikut ini :

$$x = (X - \bar{X})$$

$$y = (Y - \bar{Y})$$

$$\text{dan } xy = (X - \bar{X})(Y - \bar{Y})$$

Nilai dari a dan b pada persamaan regresi dapat dihitung dengan rumus berikut :

$$b = \frac{\sum x_i y_i}{\sum x_i^2}$$

$$b = \frac{n \sum X_i Y_i - \sum X_i \sum Y_i}{n \sum X_i^2 - (\sum X_i)^2}$$

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

Tabel perhit.

Karyawan	Hasil Produksi (lusin) (Y)	Skor Tes (X)	y (Y - \bar{Y})	x (X - \bar{X})	xy	x ²	y ²
A	30	6	-7	-1	7	1	49
B	49	9	12	2	24	4	144
C	18	3	-19	-4	76	16	361
D	42	8	5	1	5	1	25
E	39	7	2	0	0	0	4
F	25	5	-12	-2	24	4	144
G	41	8	4	1	4	1	16
H	52	10	15	3	45	9	225
	296	56	0	0	185	36	968

$$\bar{Y} = \frac{\sum Y}{N} = \frac{296}{8} = 37$$

$$\bar{X} = \frac{\sum X}{N} = \frac{56}{8} = 7$$

$$b = \frac{\sum xy}{\sum x^2} = \frac{185}{36} = 5,138 \sim 5,14$$

$$a = \bar{Y} - b\bar{X} = 37 - 5,14(7) = 1,02$$

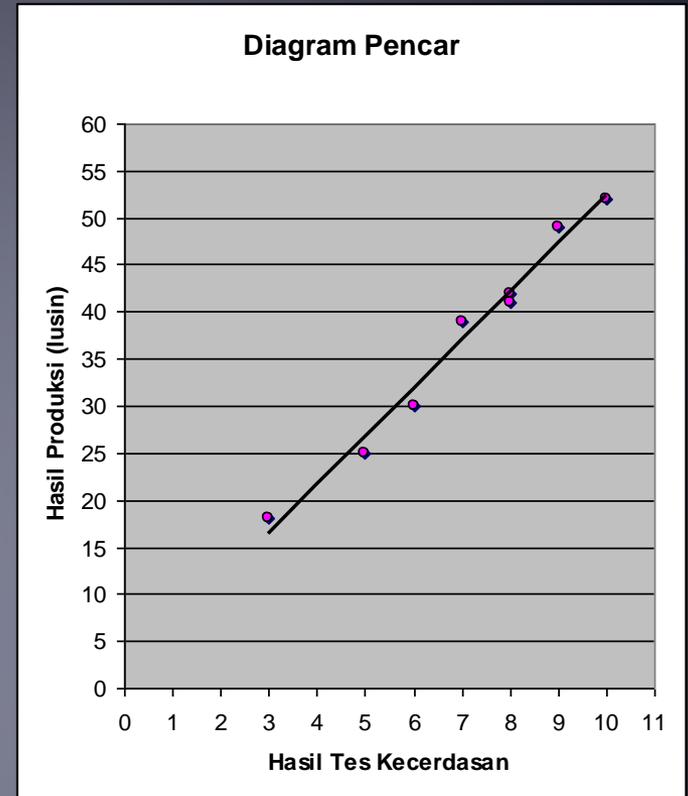
$$Y' = 1,02 + 5,14X$$

$$X = 3 \rightarrow Y' = 1,02 + 5,14(3) = 16,44$$

$$X = 5 \rightarrow Y' = 1,02 + 5,14(5) = 26,72$$

$$X = 6 \rightarrow Y' = 1,02 + 5,14(6) = 31,86$$

$$X = 10 \rightarrow Y' = 1,02 + 5,14(10) = 52,42$$



Pembacaan garis regresi

Contoh: Jika garis regresi ditunjukkan dengan persamaan :

$$y^1 = 2,94 + 0,95x$$

maka dapat diartikan bahwa : hasil penjualan akan bertambah sebesar 0,95 pada setiap kenaikan pengeluaran iklan sebanyak 1 satuan

Koefisien Regresi

- Adalah lereng garis regresi (nilai b)
- Nilai b positif, menunjukkan hubungan antara variabel x dan y searah atau hubungannya positif.
- Nilai b negatif, menunjukkan hubungan antara variabel x dan y berlawanan arah atau hubungannya negatif
- Besar kecilnya perubahan variabel x terhadap variabel y ditentukan besar kecilnya koefisien regresi.

Perbedaan Regresi dan Korelasi

- Regresi menunjukkan hubungan antara variabel satu dengan variabel lainnya.
- Sifat hubungan dapat dijelaskan: variabel yang satu sebagai penyebab, variabel yang lain sebagai akibat.
- Korelasi tidak menunjukkan hubungan sebab akibat, akan tetapi menunjukkan hubungan antara variabel satu dengan yang lain.

Interpretasi penyimpangan standar terhadap garis regresi (standart error of estimate)

- Penyimpangan standar terhadap garis regresi diinterpretasikan sebagai penyimpangan terhadap rata-rata.
- Semakin besar nilai Se semakin tersebar titik-titik yang berada di sekitar garis regresi.
- Apabila $Se=0$ atau penyimpangan standar terhadap garis regresi = 0, maka semua titik berada di sepanjang garis regresi.
- $Se=0$ maka persamaan garis regresi dapat digunakan secara sempurna untuk menaksir variabel dependen.

Koefisien Determinasi

- Adalah alat utama untuk mengetahui sejauh mana tingkat hubungan antara variabel x dan y.
- Nilai koefisien determinasi antara $0 \leq r^2 \leq 1$
- Nilai koefisien determinasi = 1 menunjukkan hubungan sempurna.
- Nilai koefisien determinasi = 0 menunjukkan tidak ada hubungan.
- $r^2 = 81$ artinya 81% perubahan dari variabel y ditentukan oleh variabel x.

Koefisien korelasi

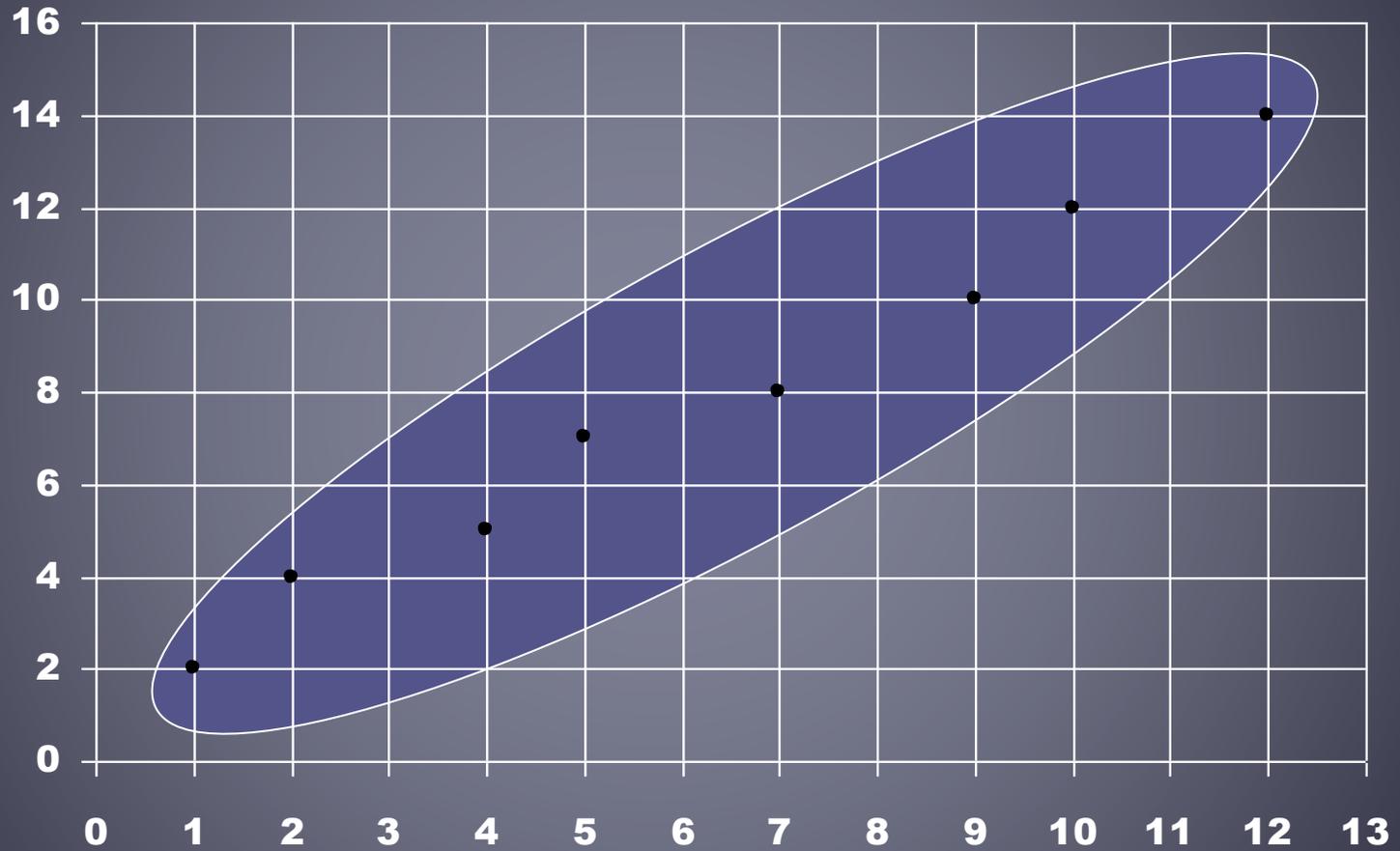
- Adalah alat kedua untuk menjelaskan hubungan antara variabel x dan y.
- Koefisien korelasi merupakan akar dari koefisien determinasi ($r = \sqrt{r^2}$)
- Koefisien korelasi menunjukkan arah hubungan antara variabel x dan y.

Hubungan dua variabel ada yang positif dan negatif.

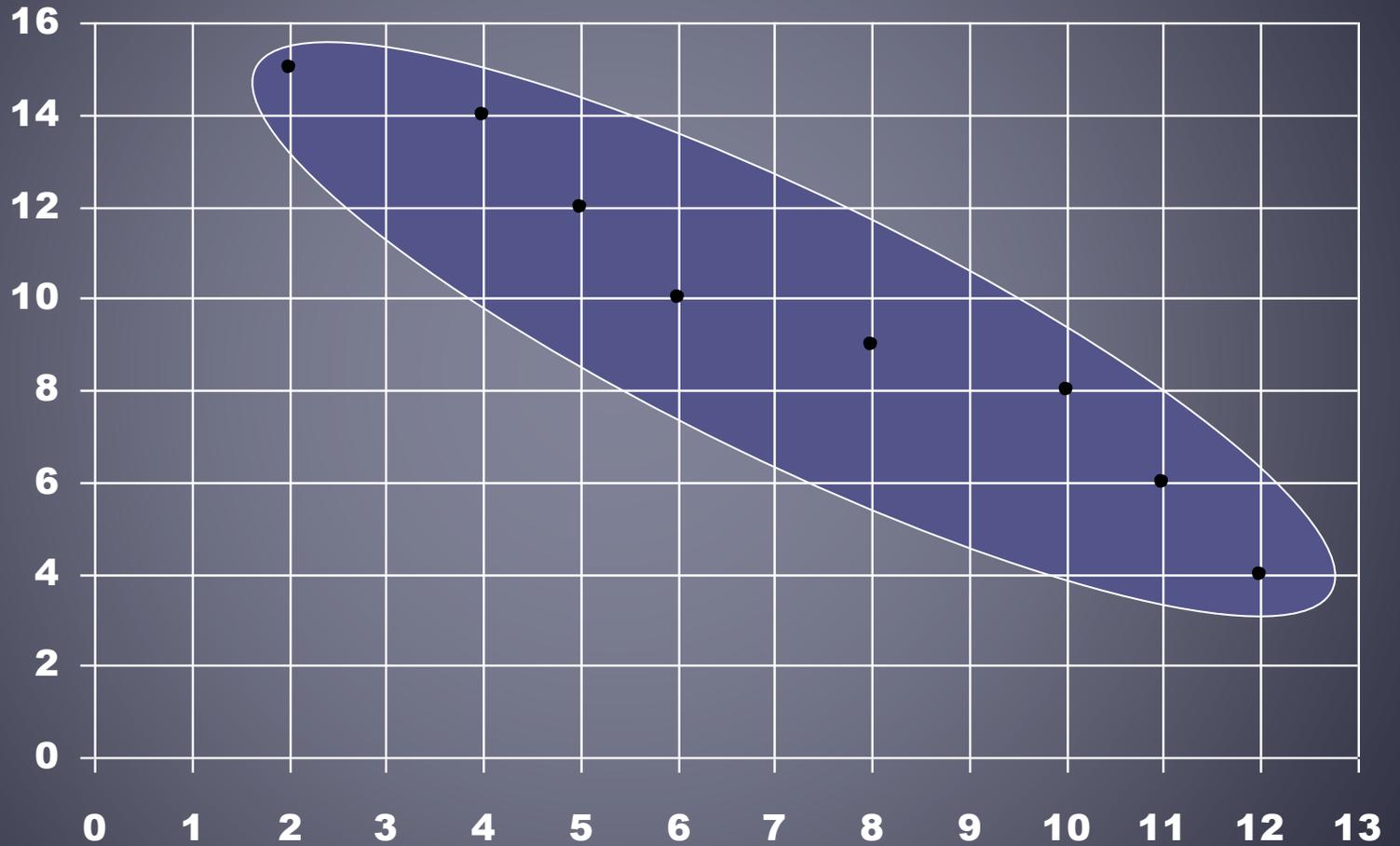
Hubungan X dan Y dikatakan positif apabila kenaikan (penurunan) X pada umumnya diikuti oleh kenaikan (penurunan) Y.

Sebaliknya dikatakan negatif kalau kenaikan (penurunan) X pada umumnya diikuti oleh penurunan (kenaikan) Y.

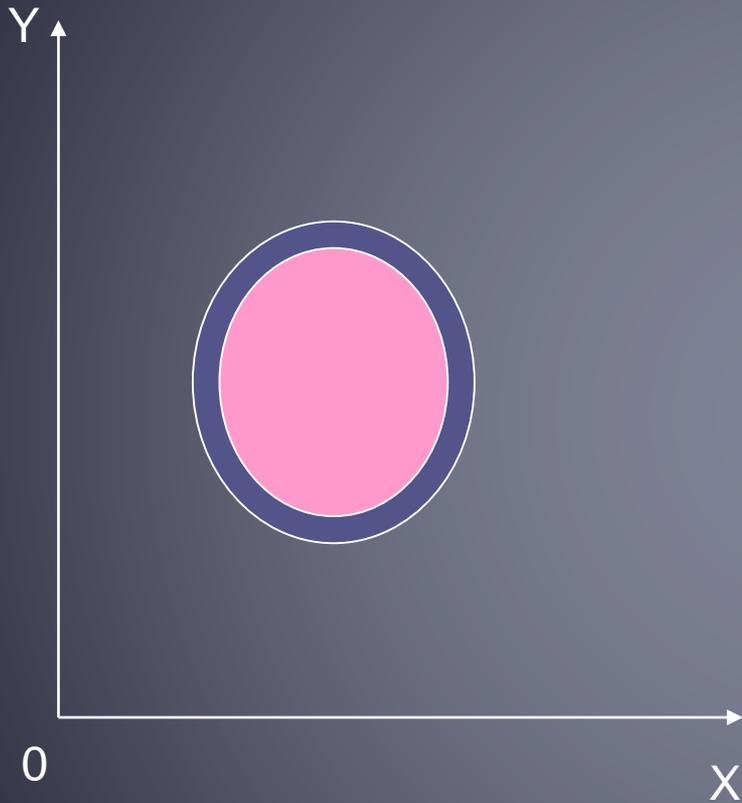
Koefisien korelasi (x dan y) mempunyai hubungan positif



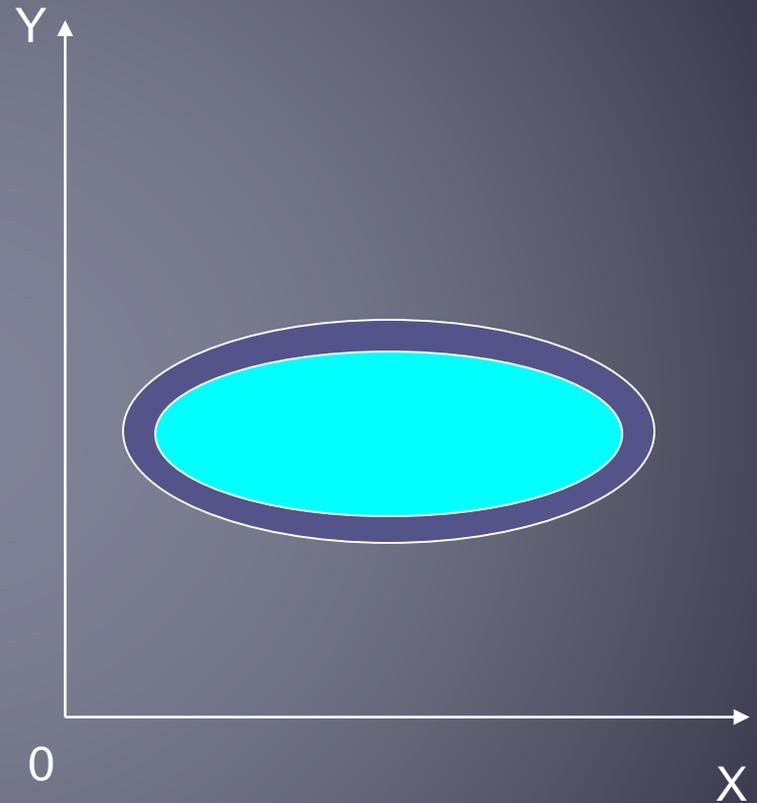
Koefisien korelasi (x dan y) mempunyai hubungan negatif



Koefisien korelasi (x dan y) tidak mempunyai hubungan atau hubungan lemah sekali



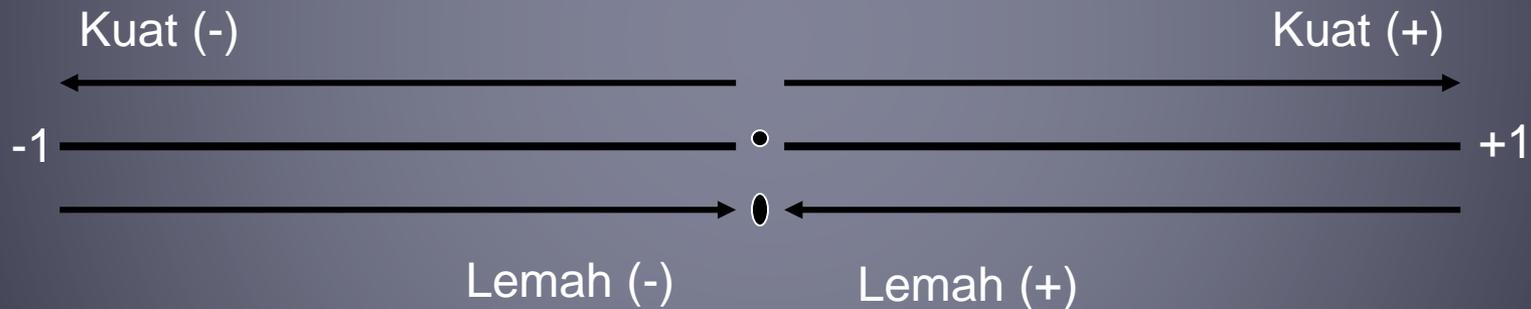
atau



Kuat dan tidaknya hubungan antara X dan Y apabila dapat dinyatakan dengan fungsi linear (paling tidak mendekati), diukur dengan suatu nilai yang disebut koefisien korelasi. Nilai koefisien korelasi ini paling sedikit -1 dan paling besar $+1$.

Jadi jika $r =$ koefisien korelasi, maka r dapat dinyatakan sebagai berikut :

$$-1 \leq r \leq +1$$



Jika $r = +1$, hubungan X dan Y sempurna dan positif,
 $r = -1$, hubungan X dan Y sempurna dan negatif,
 r mendekati $+1$, hubungan sangat kuat dan positif,
 r mendekati -1 , hubungan sangat lemah dan negatif.

Kalau koefisien penentuan ditulis KP, maka untuk menghitung KP digunakan rumus berikut : $KP = r^2$

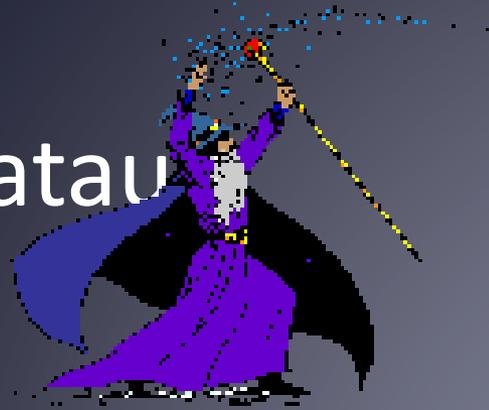
Cara menghitung r adalah sebagai berikut:

$$r = \frac{\sum_{i=1}^n x_i y_i}{\sqrt{\sum_{i=1}^n x_i^2} \sqrt{\sum_{i=1}^n y_i^2}}$$

(7.2)



atau



$$r = \frac{n \sum_{i=1}^n X_i Y_i - \sum_{i=1}^n X_i \sum_{i=1}^n Y_i}{\sqrt{n \sum_{i=1}^n X_i^2 - \left(\sum_{i=1}^n X_i \right)^2} \sqrt{n \sum_{i=1}^n Y_i^2 - \left(\sum_{i=1}^n Y_i \right)^2}}$$

Kedua rumus diatas disebut koefisien korelasi **Pearson**

Contoh 7.1

X	1	2	4	5	7	9	10	12
Y	2	4	5	7	8	10	12	14

Tabel 7.2

X	Y	$X - \bar{X}$	$Y - \bar{Y}$	x^2	y^2	xy
		(x)	(y)			
1	2	-5,25	-5,75	27,5625	33,0625	30,1875
2	4	-4,25	-3,75	18,0625	14,0625	15,9375
4	5	-2,25	-2,75	5,0625	7,5625	6,1875
5	7	-1,25	-0,75	1,5625	0,5625	0,9375
7	8	0,75	0,25	0,5625	0,0625	0,1875
9	10	2,75	2,25	7,5625	5,0625	6,1875
10	12	3,75	4,25	14,0625	18,0625	15,9375
12	14	5,75	6,25	33,0625	39,0625	35,9375
$\sum X_i = 50$	$\sum Y_i = 62$	$\sum x_i = 0$	$\sum y_i = 0$	$\sum x_i^2 = 107,5$	$\sum y_i^2 = 117,5$	$\sum x_i y_i = 111,5$
$\bar{X} = 6,25$	$\bar{Y} = 7,75$					

$$r = \frac{\sum_{i=1}^8 x_i y_i}{\sqrt{\sum_{i=1}^8 x_i^2} \sqrt{\sum_{i=1}^8 y_i^2}}$$

$$r = \frac{111,5}{\sqrt{107,5} \times \sqrt{117,5}} = 0,99$$

$$KP = r^2 = (0,99)^2 = 0,9801 \times 100\% = 98\%$$