

MODUL 6

ANALISIS KORELASI LINEAR SEDERHANA



CAPAIAN PEMBELAJARAN

1. Praktikan mampu mengimplementasikan penggunaan software R untuk analisis korelasi linear sederhana



KEBUTUHAN ALAT/BAHAN/SOFTWARE

1. Komputer
2. Software R



DASAR TEORI

A. PENDAHULUAN

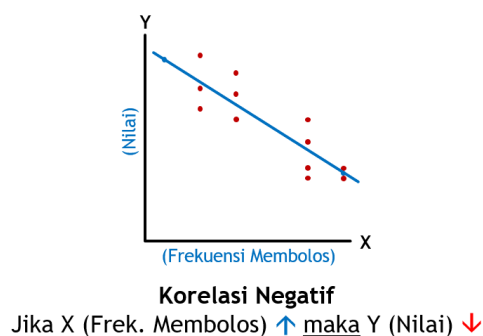
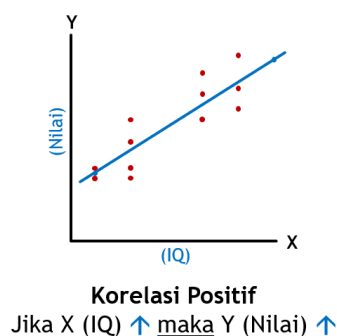
Pada bagian ini akan dibicarakan permasalahan mengenai bagaimanakah menentukan kekuatan relasi (hubungan) linear antara variabel X dan variabel Y. Kekuatan relasi (hubungan) antara X dan Y dinyatakan dengan **Koefisien Korelasi**.

Kisaran nilai koefisien korelasi yaitu $-1.0 \leq r_{xy} \leq 1.0$, dengan rincian:

Nilai r_{xy}		Keterangan
Positif	$0.5 \leq r_{xy} \leq 1.0$	Korelasi kuat dan berbanding lurus
	$0 < r_{xy} < 0.5$	Korelasi lemah dan berbanding lurus
	$r_{xy} = 0.0$	Tidak ada korelasi
Negatif	$-0.5 < r_{xy} < 0.0$	Korelasi lemah dan berbanding terbalik
	$-1.0 \leq r_{xy} \leq -0.5$	Korelasi kuat dan berbanding terbalik

Keterangan:

- Korelasi yang **berbanding lurus (positif)** menandakan bahwa jika variable pertama nilainya **naik**, maka nilai variable kedua juga **naik**. Begitu juga sebaliknya.
- Korelasi yang **berbanding terbalik (negative)** menandakan bahwa jika variable pertama nilainya **turun**, maka nilai variable kedua justru **naik**. Begitu juga sebaliknya.
- Korelasi yang bernilai **lemah (mendekati 0)** menandakan bahwa hubungan kedua variable lemah, sehingga **kurang cocok digunakan untuk prediksi**.
- Korelasi yang bernilai **kuat (mendekati 1)** menandakan bahwa hubungan kedua variable kuat, sehingga **cocok digunakan untuk prediksi**.



B. PERSAMAAN UNTUK MENCARI KOEFISIEN KORELASI

NAMA	PERSAMAAN
Karl-Pearson (Product Moment)	$r_{xy} = \frac{n(\sum XY) - (\sum X)(\sum Y)}{\sqrt{(n(\sum X^2) - (\sum X)^2)(n(\sum Y^2) - (\sum Y)^2)}}$
Spearman-Brown	$r_s = 1 - \frac{6\sum d_i^2}{n(n^2 - 1)}$
Kendall Tau	$\tau = \frac{\sum A - \sum B}{\frac{N(N-1)}{2}}$



PRAKTIK

Praktik 1 (Koefisien Korelasi Linear Pearson dengan R)

Diketahui **datakorelasi1** sebagai berikut

X1	Y1	X3	Y3
1	1	1	10
2	2	2	9
3	3	3	8
4	4	4	7
5	5	5	6
6	6	6	5
7	7	7	4
8	8	8	3
9	9	9	2
10	10	10	1

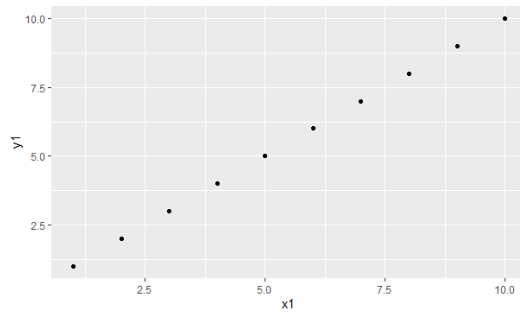
Gambarkan plot dari **datakorelasi1** dan carilah nilai koefisien korelasi pearson antara **X1** dengan **Y1**, dan **X3** dengan **Y3**!

Inputkan data dalam R (pada bagian Editor atau Console). Setelah itu, aktifkan package `ggplot2`, dengan maksud untuk menggunakan fungsi `ggplot`. Fungsi `ggplot` bertujuan untuk menyajikan grafik sebaran data.

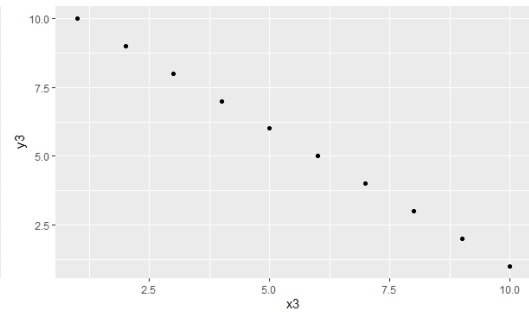
Secara umum, perintah untuk menghitung koefisien korelasi linear Pearson dalam R sebagai yaitu: `cor(variabel1, variabel2, method = "pearson")`

```
1 x1<-c(1,2,3,4,5,6,7,8,9,10)
2 y1<-c(1,2,3,4,5,6,7,8,9,10)
3 x3<-c(1,2,3,4,5,6,7,8,9,10)
4 y3<-c(10,9,8,7,6,5,4,3,2,1)
5 datakorelasi1<-data.frame(x1,y1,x3,y3)
6
7 #menggambar plot
8 library(ggplot2)
9 ggplot(datakorelasi1,aes(x1,y1))+geom_point()
10 ggplot(datakorelasi1,aes(x3,y3))+geom_point()
11
12 #mencari koefisien korelasi dengan pearson
13 cor(datakorelasi1$x1,datakorelasi1$y1,method = "pearson")
14 cor(datakorelasi1$x3,datakorelasi1$y3,method = "pearson")
15 |
```

Output



Gambar 1a



Gambar 1b

```
Console Terminal x Jobs x
> #menggambar plot
> library(ggplot2)
> ggplot(datakorelasi1,aes(x1,y1))+geom_point()
> ggplot(datakorelasi1,aes(x3,y3))+geom_point()
>
> #mencari koefisien korelasi dengan pearson
> cor(datakorelasi1$x1,datakorelasi1$y1,method = "pearson")
[1] 1
> cor(datakorelasi1$x3,datakorelasi1$y3,method = "pearson")
[1] -1
```

Gambar 1c

Analisis

Gambar 1a menunjukkan terjadinya hubungan **linear positif** yang **sempurna** antara X1 dan Y1, apabila nilai koefisien korelasi Pearson dihitung, maka akan bernilai 1 (Gambar 1c). **Hubungan positif berarti sebaran data cenderung menyebar dari kiri bawah ke kanan atas.**

Koef. Korelasi Pearson X1 dan Y1 = 1 (kuat dan berbanding lurus).

Gambar 1b menunjukkan terjadinya hubungan **linear negatif** yang **sempurna** antara X3 dan Y3, apabila nilai koefisien korelasi Pearson dihitung, maka akan bernilai -1 (Gambar 1c). **Hubungan negatif berarti sebaran data cenderung menyebar dari kiri atas ke kanan bawah.**

Koef. Korelasi Pearson X3 dan Y3 = -1 (kuat dan berbanding terbalik).

Praktik 2 (Koefisien Korelasi Linear Pearson dengan R)

Diketahui **datakorelasi2** sebagai berikut

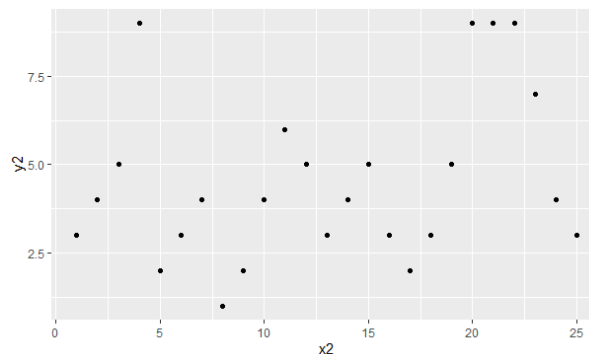
X2	Y2
1	3
2	4
3	5
4	9
5	2
6	3
7	4
8	1
9	2
10	4
11	6
12	5
13	3
14	4
15	5
16	3
17	2
18	3
19	5
20	9
21	9
22	9
23	7
24	4
25	3

Gambarkan plot dari **datakorelasi2** dan carilah nilai koefisien korelasi pearson antara **X2** dengan **Y2**!

Inputkan data dalam R (pada bagian Editor atau Console).

```
1 x2<-c(1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25)
2 y2<-c(3,4,5,9,2,3,4,1,2,4,6,5,3,4,5,3,2,3,5,9,9,9,7,4,3)
3 datakorelasi2<-data.frame(x2,y2)
4
5 #menggambar plot
6 library(ggplot2)
7 ggplot(datakorelasi2,aes(x2,y2))+geom_point()
8
9 #mencari koefisien korelasi dengan pearson
10 cor(datakorelasi2$x2,datakorelasi2$y2,method = "pearson")
11
```

Output



Gambar 2a

```
Console Terminal Jobs x
~ /
> x2<-c(1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25)
> y2<-c(3,4,5,9,2,3,4,1,2,4,6,5,3,4,5,3,2,3,5,9,9,9,7,4,3)
> datakorelasi2<-data.frame(x2,y2)
>
> #menggambar plot
> library(ggplot2)
> ggplot(datakorelasi2,aes(x2,y2))+geom_point()
>
> #mencari koefisien korelasi dengan pearson
> cor(datakorelasi2$x2,datakorelasi2$y2,method = "pearson")
[1] 0.304238
>
```

Gambar 2b

Analisis

Sebaran data pada Gambar 2a cenderung **acak (tidak beraturan)**, sehingga hubungan linear yang terjadi antara X2 dan Y2 lemah. Apabila nilai koefisien korelasi Pearson dihitung, maka akan bernilai mendekati 0 (Gambar 2b).

Koef. Korelasi Pearson X2 dan Y2 = 0 (tidak ada korelasi).

Praktik 3 (Koefisien Korelasi Linear Pearson Seluruh Data)

Secara umum, perintah untuk menghitung koefisien korelasi linear Pearson dalam R sebagai yaitu: `cor(data)`

```
Console Terminal Jobs x
~ /
> #mencari koefisien korelasi dengan pearson seluruh data
> cor(datakorelasi1)
      x1 y1 x3 y3
x1  1  1  1 -1
y1  1  1  1 -1
x3  1  1  1 -1
y3 -1 -1 -1  1
```

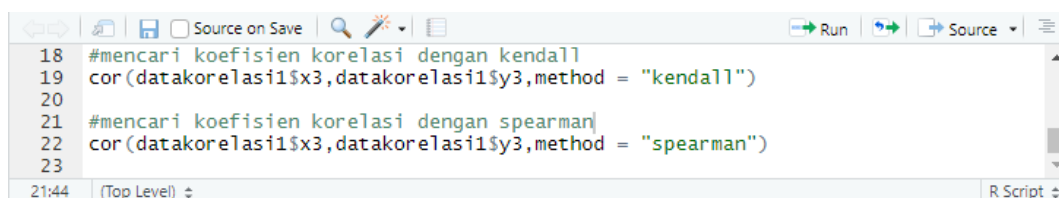
Gambar 3

Analisis

Pada Gambar 3 terlihat koefisien korelasi linear Pearson dalam bentuk tabel tabulasi silang (menurut kolom dan baris). Dapat dilihat, misalkan korelasi antara X1 (baris 1) dengan Y3 (kolom 3) sebesar -1 (korelasi negative sempurna). Dengan menggunakan perintah `cor(data)`, dapat digunakan untuk menampilkan nilai koefisien korelasi Pearson seluruh pasangan variable pada suatu data.

Praktik 4 (Koefisien Korelasi Linear Kendall dan Spearman dengan R)

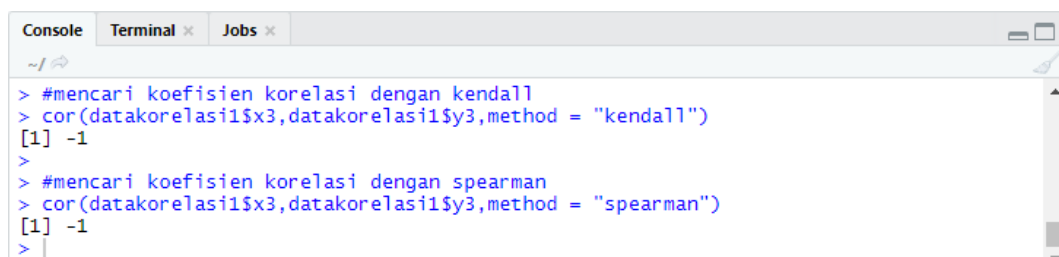
Selain menggunakan Pearson (yang lazim digunakan), pencarian koefisien korelasi linear juga dapat menggunakan persamaan Kendall dan Spearman. Implementasi dalam R yaitu:



```
18 #mencari koefisien korelasi dengan kendall
19 cor(datakorelasi1$x3,datakorelasi1$y3,method = "kendall")
20
21 #mencari koefisien korelasi dengan spearman
22 cor(datakorelasi1$x3,datakorelasi1$y3,method = "spearman")
23
```

Gambar 4a

Output



```
> #mencari koefisien korelasi dengan kendall
> cor(datakorelasi1$x3,datakorelasi1$y3,method = "kendall")
[1] -1
>
> #mencari koefisien korelasi dengan spearman
> cor(datakorelasi1$x3,datakorelasi1$y3,method = "spearman")
[1] -1
>
```

Gambar 4b

Analisis

Gambar 4b menunjukkan output pencarian korelasi dengan kendall dan spearman antara X3 dan Y3.

Koef. Korelasi Kendall X3 dan Y3 = -1 (kuat dan berbanding terbalik).

Koef. Korelasi Spearman X3 dan Y3 = -1 (kuat dan berbanding terbalik).



LATIHAN

A. Data Nilai UAS Mata Kuliah Matematika dan Bahasa Inggris untuk 12 mahasiswa adalah sebagai berikut:

Matematika (X)	Bahasa Inggris (Y)
60	80
45	69
50	71
60	85
50	80
65	82
60	89
65	93
50	76
65	86
45	71
50	69

1. Gambarlah plot/diagram pencar dari data tersebut!
2. Carilah koefisien korelasinya menggunakan persamaan Pearson, Kendall dan Spearman, serta tentukan makna koefisien korelasi yang diperoleh!

B. Data Nilai UTS Mata Kuliah Statistika (X) dan Statistika Terapan (Y) untuk 12 mahasiswa adalah sebagai berikut:

X	Y
6	7
7	8
5	6
9	9
8	7
7	6
5	6
6	7
4	5
3	4
8	8
7	7

1. Gambarlah plot/diagram pencar dari data tersebut!
2. Carilah koefisien korelasinya menggunakan persamaan Pearson, Kendall dan Spearman, serta tentukan makna koefisien korelasi yang diperoleh!



TUGAS

Data nilai-nilai ujian masuk (X) dan rata-rata nilai pada tahun pertama (Y) dari 20 mahasiswa.

X	Y
55	61
79	72
59	69
81	89
62	72
70	75
80	61
89	79
92	90
60	55
63	85
64	87
69	70
75	90
84	67
77	84
62	72
85	70
55	60
66	67

1. Gambarlah plot/diagram pencar dari data tersebut!
2. Carilah koefisien korelasinya menggunakan persamaan Pearson, Kendall dan Spearman, serta tentukan makna koefisien korelasi yang diperoleh!



REFERENSI

- [1] Gio, P.U., Effendie, A.R. 2017. Belajar Bahasan Pemrograman R (Dilengkapi Cara Membuat Aplikasi Olah Data Sederhana dengan R Shiny). Medan: USU Press.