R dalam Analisis korelasi dan Regresi linier sederhana

BAB 2

Input Data – Cara I

IQ	IPK
99	3.00
118	3.50
108	3.20
110	3.25
97	2.70
120	3.50
105	3.25

- IQ <- c(99, 118, 108, 110, 97, 120, 105)
- IPK <-c(3, 3.5, 3.2, 3.25, 2.7, 3.5, 3.25)

Input Data – Cara I

IQ	IPK
99	3.00
118	3.50
108	3.20
110	3.25
97	2.70
120	3.50
105	3.25

- IQ <- scan()

  1 : 99 118 108 110 97 120 105

  8: (*enter*)

- IPK <-scan()

  1: 3 3.5 3.2 3.25 2.7 3.5 3.25

Example

```
> IQ<-scan()
1: 99
2: 118
3: 108
4: 110
5: 97
6: 120
7: 105
8:
Read 7 items
> IPK<-scan()
1: 3
2: 3.5
3: 3.2
4: 3.25
5: 2.7
6: 3.5
7: 3.25
8:
Read 7 items
```

Analisis Korelasi

- Mengetahui nilai korelasi
- Melakukan uji korelasi

IQ	IPK
99	3.00
118	3.50
108	3.20
110	3.25
97	2.70
120	3.50
105	3.25

Menggunakan command

> cor(IPK,IQ)

> cor.test(IPK,IQ)

Output

1. Nilai korelasi

```
> cor(IPK,IQ)
[1] 0.9381326
> cor.test(IPK,IQ)

        Pearson's product-moment correlation

data:  IPK and IQ
t = 6.058, df = 5, p-value = 0.001768
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.6305094 0.9910472
sample estimates:
      cor
0.9381326
```

- Nilai Korelasi antara IPK, IQ adalah 0.9381326
- Kesimpulan Uji Hipotesis Nilai Korelasi tidak sama dengan 0

2. Melakukan uji korelasi

```
> cor(IPK,IQ)
[1] 0.9381326
> cor.test(IPK,IQ)

        Pearson's product-moment correlation

data:  IPK and IQ
t = 6.058, df = 5, p-value = 0.001768
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.6305094 0.9910472
sample estimates:
      cor
0.9381326
```

i. Menyusun hipotesis

H0 : Tidak ada korelasi antara IPK dan IQ (cor(IPK,IQ) = 0)

H1 : Terdapat korelasi antara IPK dan IQ (cor(IPK, IQ) <> 0)

```
> cor(IPK,IQ)
[1] 0.9381326
> cor.test(IPK,IQ)

        Pearson's product-moment correlation

data:  IPK and IQ
t = 6.058, df = 5, p-value = 0.001768
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.6305094 0.9910472
sample estimates:
      cor
0.9381326
```

ii. Pilih tingkat signifikansi α=5%=0.05

iii. Daerah Kritis (daerah penolakan hipotesis null)

- P-value < α = 0.05
- Karena α = 0.05 > p-value=0.001768 maka H0 ditolak
- D.k.l ada korelasi antara IPK dan IQ

Analisis Regresi

- Membuat persamaan
- Mengetahui hasil regresi
- Mengetahui uji anova

```
> fm <- lm(IPK ~ IQ)
> summary(fm)
> anova(fm)
```

Persamaannya?

IPK = -0.064150 + 0.030184*IQ

```
> fm<-lm(IPK~IQ)
> summary(fm)

Call:
lm(formula = IPK ~ IQ)

Residuals:
       1        2        3        4        5        6        7
 0.075965 0.002475 0.004312 -0.006055 -0.163667 -0.057892 0.144863

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.064150   0.540328  -0.119  0.91012
IQ           0.030184   0.004982   6.058  0.00177 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1067 on 5 degrees of freedom
Multiple R-squared:  0.8801,    Adjusted R-squared:  0.8561
F-statistic:  36.7 on 1 and 5 DF,  p-value: 0.001768
```

Uji Hipotesisnya

```
> anova(fm)
Analysis of Variance Table

Response: IPK
          Df  Sum Sq Mean Sq F value   Pr(>F)
IQ         1 0.41804 0.41804  36.699 0.001768 **
Residuals  5 0.05696 0.01139
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- Hipotesis
  - H-null : B = 0
  - H-alt : B ≠ 0
- Daerah Kritis: p < 0.05
- Statistik uji: F = 36.699; p=0.001768
- Hasil uji

  p = 0.001768 < $\alpha = 0.05$
- Kesimpulan

  Hipotesis null tidak diterima, dkl. IQ mempunyai pengaruh terhadap IPK

Regresi pada R

```
> fm<-lm(IPK~IQ)
> summary(fm)

Call:
lm(formula = IPK ~ IQ)

Residuals:
       1        2        3        4        5        6        7
 0.075965 0.002475 0.004312 -0.006055 -0.163667 -0.057892 0.144863

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.064150   0.540328  -0.119  0.91012
IQ           0.030184   0.004982   6.058  0.00177 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1067 on 5 degrees of freedom
Multiple R-squared:  0.8801,    Adjusted R-squared:  0.8561
F-statistic:  36.7 on 1 and 5 DF,  p-value: 0.001768
```

Uji ANOVA

```
> anova(fm)
Analysis of Variance Table

Response: IPK
          Df  Sum Sq Mean Sq F value   Pr(>F)
IQ         1 0.41804 0.41804  36.699 0.001768 **
Residuals  5 0.05696 0.01139
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Uji kenormalan

- Dengan menggunakan uji shapiro wilk

```
> shapiro.test(residuals(fm))

        Shapiro-Wilk normality test

data:  residuals(fm)
W = 0.96361, p-value = 0.8491
```

- i. Menyusun hipotesis
  - H0: residual berdistribusi normal
  - H1: residual tidak berdistribusi normal

ii. Pilih tingkat signifikansi α=0.05

iii. Statistika Uji Shapiro Wilk

→ W=0.96361

→ p-value=0.8491

→ Karena α=0.05< p-value=0.8491 maka H0 tidak ditolak

→ D.k.l asumsi kenormalan dapat dipenuhi

Mencari residual

- Menghitug residual terstandar dengan fungsi rsstandard

```
> sres<-rstandard(fm)
> sres[1:5]
         1          2          3          4          5
 0.86633904 0.02886657 0.04363896 -0.06155248 -2.00228192
> sres[1:7]
         1          2          3          4          5          6
 0.86633904 0.02886657 0.04363896 -0.06155248 -2.00228192 -0.73090337
         7
 1.48480525
```

- Mencari observasi yang diduga outlier (studentized residual)

```
> sres[which(abs(sres)>2)]
        5
-2.002282
```