Praktikum Eksplorasi dan Visualisasi Data
Pertemuan 9
Analisis Regresi Linear Konfirmasi

A. Pengertian

Istilah regresi pertama kali diperkenalkan pada tahun 1886 oleh Sir Francis Galton dalam penelitian biogenesisnya, yaitu mengenai hubungan antara tinggi badan anak dengan tinggi badan orang tuanya. Awal abad XIX, Carl Friedrich Gauss mempopulerkan metode kuadrat terkecil (Least Squares Methods) yang menjadi dasar analisis regresi (dalam estimasi parameter regresi).

Analisis regresi merupakan suatu analisis yang terjadi di antara dua variabel, yaitu variabel independen (prediktor) dan variabel dependen (respon) dimana variabel prediktor diasumsikan mempengaruhi variabel respon secara linear, sehingga variabel respon dapat diduga dari variabel prediktor. Linear disini artinya bila pasangan data observasi digambarkan pada suatu bidang kartesius dengan variabel independen sebagai sumbu X dan variabel dependen sebagai sumbu Y, maka sebaran titik-titiknya mengikuti garis lurus.

B. Tujuan

Tujuan dari analisis regresi antara lain, sebagai berikut:

- Mengetahui pola keeratan hubungan
- Mengestimasi pengaruh variabel prediktor terhadap variabel respon
- Memprediksi besarnya variabel respon jika besarnya prediktor diketahui

C. Model Regresi Linear

Analisis regresi linear sederhana adalah analisis regresi yang terjadi antara variabel dependen dengan satu variabel independen dimana variabel independen diasumsikan mempengaruhi variabel dependen secara linear. Model populasi dari regresi linear sederhana adalah

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$

$$E(Y_i) = \beta_0 + \beta_1 X_i$$

Dengan:

- Y_i : variabel dependen pada observasi ke-i
- β_0 : parameter konstanta regresi (merupakan nilai Y bila variabel independen = 0 (nilai konstan))
- β_1 : parameter koefisien regresi (menunjukkan angka peningkatan atau penurunan variabel dependen yang didasarkan pada variabel independen)
- X_i : variabel independen pada observasi ke-i
- ε_i : nilai random error dengan rata-rata $E\{\varepsilon_i\} = 0$ dan variansi nya $Var\{\varepsilon_i\} = \sigma^2$; ε_i dan ε_j tidak berkorelasi sehingga nilai kovariansinya $Cov\{\varepsilon i, \varepsilon j\} = 0$ untuk semua nilai i dan j, $i \neq j$, i,j = 1, 2, 3, …,n.

β_0 dan β_1 merupakan parameter yang tidak diketahui nilainya dan akan diestimasi dengan statistik b0 dan b1 dengan menggunakan metode kuadrat terkecil (Least Squares Method). Hasil dari penurunan estimator koefisien regresi yang berturut-turut sebagai berikut:

$$b1 = \frac{n\sum_i X_i Y_i - \sum_i X_i \sum_i Y_i}{n\ \sum_i X_i^2 - (\sum_i X_i)^2}$$

$$b0 = Y - b1X$$

D. Analisis korelasi

Korelasi digunakan untuk mencari hubungan dan membuktikan hipotesis hubungan dua variabel bila data kedua variabel berskala interval atau rasio. Koefisien korelasi untuk populasi diberi simbol rho (ρ) dan untuk sampel diberi simbol r, sedangkan untuk korelasi ganda diberi simbol R. Rumus yang dapat digunakan untuk menghitung koefisien korelasi sebagai berikut

$$r_{xy} = \frac{n\sum x_i y_i - (\sum x_i)(\sum y_i)}{\sqrt{\{n\sum x_i^2 - (\sum x_i)^2\} - \{n\sum y_i^2 - (\sum y_i)^2\}}}$$

Penafsiran terhadap koefisien korelasi untuk menentukan besar atau kecilnya korelasi dapat berpedoman pada ketentuan yang tertera pada tabel di bawah ini:

Interval Koefisien	Tingkat Hubungan
0,00 – 0,199	Sangat rendah
0,20 – 0,399	Rendah
0,40 – 0,599	Sedang
0,60 – 0,799	Kuat
0,80 – 1,00	Sangat Kuat

Interpretasi terhadap koefisien korelasi sebagai berikut :

- Nilai r berada diantara -1 sampai 1
- Nilai r+ berarti X berpengaruh secara positif terhadap Y. Semakin dekat ke 1, semakin sempurna hubungan linear X dan Y (Semakin besar nilai X, maka nilai Y juga semakin besar),
- Nilai r- berarti X berpengaruh secara berlawanan terhadap Y. Semakin dekat ke -1, semakin sempurna hubungan linear terbalik X dan Y (Semakin besar nilai X, maka nilai Y akan semakin kecil),
- Nilai r mendekati 0 berarti X tidak berpengaruh secara linear terhadap Y.

Dalam analisis korelasi terdapat suatu angka yang disebut dengan koefisien determinasi yang besarnya adalah kuadrat dari koefisien korelasi (r^2). Koefisien determinasi disebut penentu karena varian yang terjadi pada variabel dependen dapat dijelaskan melalui varian yang terjadi pada variabel independen.

E. Alur Analisis Regresi Linear Sederhana

Uji Asumsi

- **Variabel dependen merupakan variabel kontinu**
- **Linearitas**
  Linearitas dapat diartikan adanya hubungan linear antara variabel dependen dengan variabel independen. Pengecekan asumsi ini dapat dilihat secara eksploratif dengan menggunakan scatter plot antara variabel dependen dan variabel independen.
- **Variabel dependen berdistribusi normal**
  Sebelum dilakukannya Analisis Regresi Linear, diperlukan dahulu Asumsi Normalitas. Data yang diperoleh perlu diuji dulu apakah data tersebut berasal dari populasi yang berdistribusi normal atau tidak. Hal ini akan berpengaruh terhadap uji hipotesis yang akan digunakan selanjutnya. Pada praktikum Analisis Data Eksploratif ini, digunakan boxplot untuk melakukan uji asumsi normalitas. Sehingga, jika boxplot yang dihasilkan tidak memenuhi asumsi normalitas (tidak berdistribusi normal/mendekati), maka perlu dilakukan transformasi terhadap data terlebih dahulu.

Analisis Regresi Linear Sederhana

Setelah asumsi regresi linear sederhana dipenuhi, maka prosedur uji hipotesis yang dilakukan adalah Model Summary, Uji Overall/Simultan (Kelayakan Model Regresi), Uji Parsial untuk Konstanta, dan Uji Parsial untuk Koefisien, Model yang terbentuk.

Format penulisan :

1. **Model Summary :**
   - R = menunjukkan derajat hubungan antara variabel prediktor (independen) dan variabel respons (dependen) yaitu sebesar ..... Nilai ini menunjukkan bahwa terdapat hubungan yang sangat erat antara variabel prediktor dan variabel respons.
   - R-Square = ... menunjukkan koefisien determinasi sederhana. R2 sebesar ... % menunjukkan bahwa ..... % variasi dalam variabel dependen dapat dijelaskan oleh variabel independen .Sedangkan sisanya dijelaskan oleh variabel-variabel yang lain.
   - Adjusted Square = ... menyatakan besarnya koreksi terhadap R2 sebesar .....%.
   - Std. Error of the Estimate = ... merupakan ukuran besarnya variansi dalam model regresi adalah ....
2. **Uji Overall :**
   - Hipotesis
     H_0 : semua β_i = 0, i = 0,1,2 (model regresi tidak layak digunakan)
     H_1 : ada $\beta_i \neq 0$, i = 0,1 (model regresi layak digunakan)
     dengan :
     β_0= parameter konstanta regresi
     β_1 = parameter koefisien regresi
   - Tingkat Signifikansi

α = 0.05

- Statistik Uji
  P-Value = …
- Daerah Kritik
  H_0 ditolak jika P-Value $< \alpha$
- Kesimpulan
  Karena nilai P-Value = …. (</>) 0,05 = α, maka H_0 (ditolak / diterima) Sehingga dapat disimpulkan … (kesimpulan harus memberikan jawaban dari persoalan).
- Interpretasi
  Berdasarkan pengujian yang telah dilakukan dengan *software R.* dilakukan uji …(nama uji) dengan data … (nama datanya), ingin diuji apakah … (persoalan), dengan hipotesis H_0 : ... (kalimat isi H0). Dengan tingkat signifikansi (α) sebesar … (nilai alpha). Didapatkan hasil P-Value = …. (</>) 0,05 = α, maka H_0 (ditolak / diterima) Sehingga dapat disimpulkan … (kesimpulan harus memberikan jawaban dari persoalan).

3. **Uji Parsial :**
   Uji Parsial Konstanta

- Hipotesis
  H_0 : $\beta_0 = 0$, (konstanta regresi tidak signifikan terhadap model regresi)
  H_1 : $\beta_0 \neq 0$, (konstanta regresi signifikan terhadap model regresi)
  dengan :
  β_0= parameter konstanta regresi
- Tingkat Signifikansi
  α = 0.05
- Statistik Uji
  P-Value = …
- Daerah Kritik
  H_0 ditolak jika P-Value $< \alpha$
- Kesimpulan
  Karena nilai P-Value = …. (</>) 0,05 = α, maka H_0 (ditolak / diterima) Sehingga dapat disimpulkan … (kesimpulan harus memberikan jawaban dari persoalan).
- Interpretasi
  Berdasarkan pengujian yang telah dilakukan dengan *software R.* dilakukan uji …(nama uji) dengan data … (nama datanya), ingin diuji apakah … (persoalan), dengan hipotesis H_0 : ... (kalimat isi H0). Dengan tingkat signifikansi (α) sebesar … (nilai alpha). Didapatkan hasil P-Value = …. (</>) 0,05 = α, maka H_0 (ditolak / diterima) Sehingga dapat disimpulkan … (kesimpulan harus memberikan jawaban dari persoalan).

Uji Parsial Koefisien

- Hipotesis

H_0 : β_1 = 0, (Koefisien regresi tidak signifikan terhadap model regresi)
H_1 : $\beta_1 \neq 0$, (Koefisien regresi signifikan terhadap model regresi)
dengan :
β_1= parameter koefisien variabel …..

- Tingkat Signifikansi
  α = 0.05
- Statistik Uji
  P-Value = …
- Daerah Kritik
  H_0 ditolak jika P-Value $< \alpha$
- Kesimpulan
  Karena nilai P-Value = …. (</>) 0,05 = α, maka H_0 (ditolak / diterima) Sehingga dapat disimpulkan … (kesimpulan harus memberikan jawaban dari persoalan).
- Interpretasi
  Berdasarkan pengujian yang telah dilakukan dengan *software R.* dilakukan uji …(nama uji) dengan data … (nama datanya), ingin diuji apakah … (persoalan), dengan hipotesis H_0 : ... (kalimat isi H0). Dengan tingkat signifikansi (α) sebesar … (nilai alpha). Didapatkan hasil P-Value = …. (</>) 0,05 = α, maka H_0 (ditolak / diterima) Sehingga dapat disimpulkan … (kesimpulan harus memberikan jawaban dari persoalan).

4. **Model Regresi :**

$$\widehat{Y} = \widehat{\beta 0} + \widehat{\beta 1}$$

Interpretasi :

- setiap penambahan/ pengurangan 1 satuan *x* maka nilai variabel *y* akan bertambah/ berkurang sebesar β dengan asumsi variabel lain konstan.
- Jika variabel independen bernilai nol, maka nilai variabel dependen adalah … (konstanta).

***Notes** : Jika data yang diregresi adalah data hasil transformasi, maka model ‘dikembalikan’ ke data asli.
Contoh :
x = umur, y = tinggi badan, dan y’ = tinggi badan dengan transformasi y’ = $1/y$
didapat model regresinya adalah y’ = β0 (konstan) + β1*umur
maka model regresinya diubah menjadi :

$$y' = \beta 0 \text{ (konstan)} + \beta 1 * \text{umur}$$

$$1/y = \beta 0 \text{ (konstan)} + \beta 1 * \text{umur}$$

$$y = \frac{1}{\beta 0\ (konstan) + \beta 1 * umur}$$

Interpretasi Akhir :
Menjelaskan alur pengujian dari uji asumsi sampai model regresi yang diperoleh .

F. Contoh Soal

Seorang teknisi ingin mengetahui hubungan antara suhu ruangan dengan jumlah cacat produksi dari suatu mesin yang diakibatkannya, agar nantinya dapat memprediksi jumlah cacat produksi berdasarkan suhu ruangannya. Untuk mengetahuinya diambil data jumlah cacat produksi selama 30 hari berturut-turut dan suhu ruangan pada hari tersebut. Bantulah teknisi tersebut dalam menyusun model yang sesuai berdasarkan data yang diberi.

Penyelesaian:

Sebelum masuk ke dalam analisis kita harus melakukan uji asumsi terlebih dahulu untuk melihat apakah data yang dimiliki memenuhi asumsi yang diperlukan untuk melakukan analisis regresi.

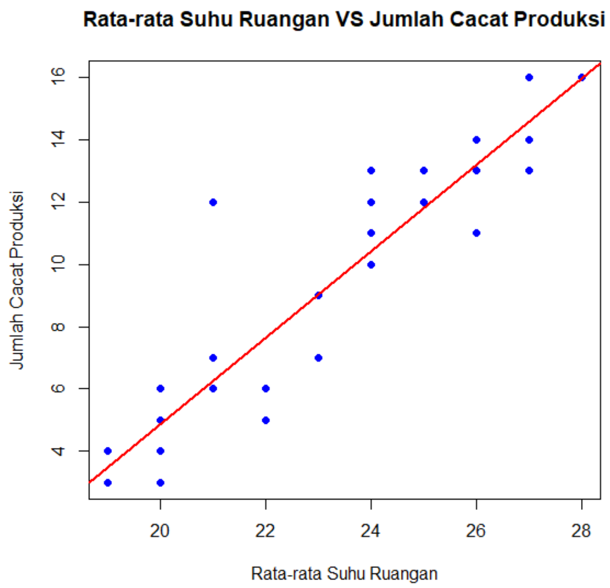
Uji Asumsi

1. **Uji Linearitas**

   Syntax :

```
# Scatterplot
plot(data$`Rata-rata Suhu Ruangan`,data$`Jumlah Cacat Produksi`,
    main="Rata-rata Suhu Ruangan VS Jumlah Cacat Produksi",
    xlab = "Rata-rata Suhu Ruangan",
    ylab = "Jumlah Cacat Produksi",col="blue",pch = 19)
# Garis Regresi linear
abline(lm(data$`Jumlah Cacat Produksi`~data$`Rata-rata Suhu Ruangan`,data = data),
      col="red",lwd=2)
```

   Output :



   Interpretasi :

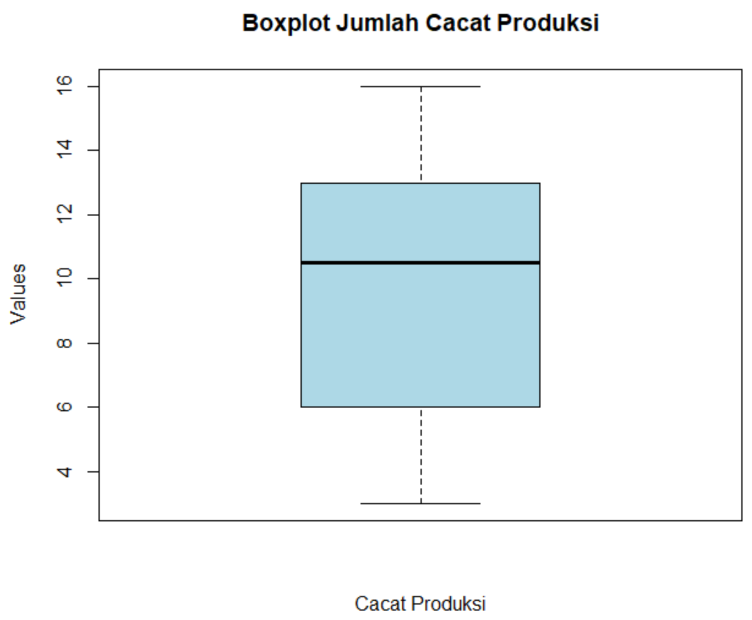
Dari grafik scatterplot yang didapatkan dapat dilihat bahwa terdapat hubungan linear antara rata-rata suhu ruangan dengan jumlah cacat produksi, dimana hubungan ini bersifat positif. Artinya adalah jika rata-rata suhu ruangan meningkat maka, jumlah cacat produksi juga meningkat.

2. **Uji Normalitas**

Syntax :

```
# Library
library(readxl)
library(car)

# Read data
data<-read_excel(file.choose(),sheet = 1)

# Uji Normalitas
boxplot(data$`Jumlah Cacat Produksi`,main="Boxplot Jumlah Cacat Produksi",
    xlab="Cacat Produksi",ylab="Values",col = "lightblue")
```

Output :



Interpretasi :

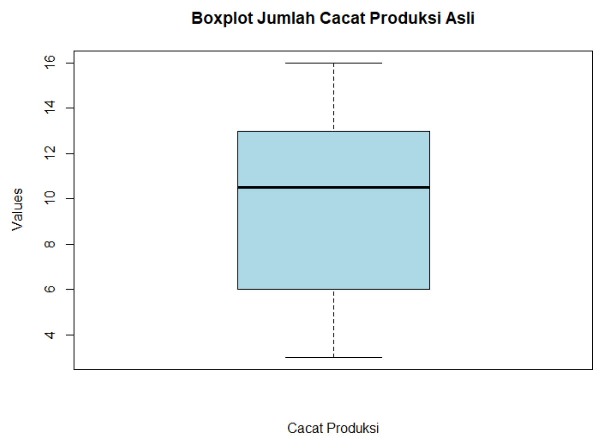
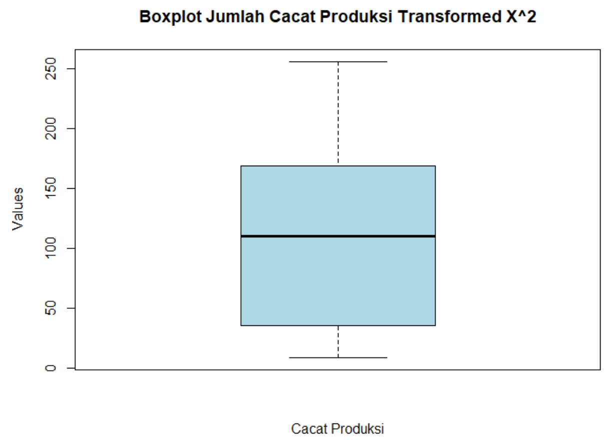
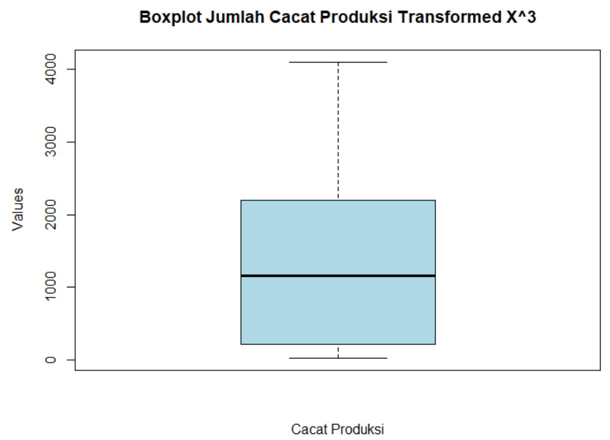
Berdasarkan output boxplot diatas yang diperoleh dengan menggunakan software R, dengan menggunakan data jumlah cacat produksi selama 30 hari berturut-turut dan suhu ruangan. Didapatkan informasi bahwa jumlah cacat produksi selama 30 hari berturut-turut diambil dari populasi yang tidak berdistribusi normal hal ini dapat diketahui dengan garis horizontal tengah (median) lebih mendekati garis horizontal atas (Q3) dari pada garis horizontal bawah (Q1). Atau dapat disebut dengan jarak antara median ke Q3 lebih kecil dari pada jarak antara median ke Q1. Jumlah cacat produksi selama 30 hari berturut-turut diambil dari populasi yang tidak berdistribusi normal, Oleh karena itu, kita akan melakukan transformasi satu angkatan untuk memperbaiki bentuk boxplot agar asumsi normalitas dapat terpenuhi.

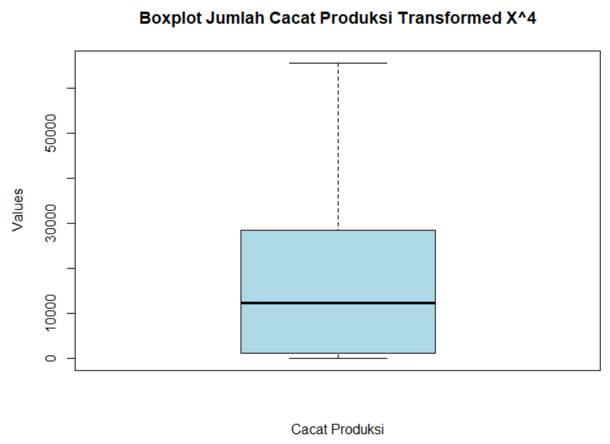
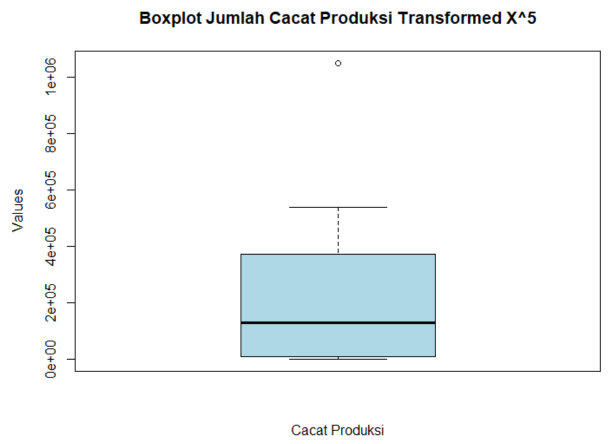
Syntax :

```
# Cek Nisbah
Nisbah_Asli = (median(data$`Jumlah Cacat Produksi`)-quantile(data$`Jumlah Cacat Produksi`,0.25))/IQR(data$`Jumlah Cacat Produksi`)
Nisbah_Asli

# Transformasi Satu angkatan
# Transformasi x^2
data$Ytf2 <-(data$`Jumlah Cacat Produksi`)^2
Nisbah_Ytf2 = (median(data$Ytf2)-quantile(data$Ytf2,0.25))/IQR(data$Ytf2)
Nisbah_Ytf2

# Transformasi x^3
data$Ytf3 <-(data$`Jumlah Cacat Produksi`)^3
Nisbah_Ytf3 = (median(data$Ytf3)-quantile(data$Ytf3,0.25))/IQR(data$Ytf3)
Nisbah_Ytf3

# Transformasi x^4
data$Ytf4 <-(data$`Jumlah Cacat Produksi`)^4
Nisbah_Ytf4 = (median(data$Ytf4)-quantile(data$Ytf4,0.25))/IQR(data$Ytf4)
Nisbah_Ytf4

# Transformasi x^5
data$Ytf5 <-(data$`Jumlah Cacat Produksi`)^5
Nisbah_Ytf5 = (median(data$Ytf5)-quantile(data$Ytf5,0.25))/IQR(data$Ytf5)
Nisbah_Ytf5


# Cek hasil transformasi
boxplot(data$Ytf2,main="Boxplot Jumlah Cacat Produksi Transformed X^2"
    ,xlab="Cacat Produksi",ylab="Values",col = "lightblue")

boxplot(data$Ytf3,main="Boxplot Jumlah Cacat Produksi Transformed X^3"
    ,xlab="Cacat Produksi",ylab="Values",col = "lightblue")

boxplot(data$Ytf4,main="Boxplot Jumlah Cacat Produksi Transformed X^4"
    ,xlab="Cacat Produksi",ylab="Values",col = "lightblue")
```

```
boxplot(data$Ytf5,main="Boxplot Jumlah Cacat Produksi Transformed X^5"
    ,xlab="Cacat Produksi",ylab="Values",col = "lightblue")
```

Output :

Transformasi	Nisbah	Outlier	Boxplot
Asli	0.6666667	Tidak ada	
x^2	0.5877712	Tidak ada	
x^3	0.5094567	Tidak ada	

x^4	0.4356003	Tidak ada	
x^5	0.3687259	Ada	

Dapat dilihat dari beberapa transformasi coba-coba yang dilakukan didapatkan transformasi x^3 paling ideal dengan nisbah yang paling mendekati 0.5 sehingga data hasil transformasi inilah yang akan digunakan dalam analisis regresi nantinya.

Analisis Regresi Linear Sederhana

1. **Model Summary**

Syntax :

```
# model summary
model <-lm(data$`Jumlah Cacat Produksi`~data$`Rata-rata Suhu Ruangan`,data = data)
glance(model)[c(1,2,3)]
cor(data$`Rata-rata Suhu Ruangan`,data$`Jumlah Cacat Produksi`)
```

Output :

```
> glance(model)[c(1,2,3)]
# A tibble: 1 x 3
  r.squared adj.r.squared sigma
      <dbl>         <dbl> <dbl>
1     0.821         0.815  1.71
> cor(data$`Rata-rata Suhu Ruangan`,data$`Jumlah Cacat Produksi`)
[1] 0.90636364
```

Untuk model summary akan digunakan yang berdasarkan data sebelum dilakukan transformasi untuk mencegah terganggunya hubungan antara variabel dependen dan variabel independen

- R = menunjukkan derajat hubungan antara variabel prediktor (independen) dan variabel respons (dependen) yaitu sebesar 0.90636364 Nilai ini menunjukkan bahwa terdapat hubungan yang sangat erat antara variabel prediktor dan variabel respons.
- R-Square = 0.821 menunjukkan koefisien determinasi sederhana. R^2 sebesar 82.1 % menunjukkan bahwa 82.1 % variasi dalam variabel dependen dapat dijelaskan oleh variabel independen .Sedangkan sisanya dijelaskan oleh variabel-variabel yang lain.
- Adjusted Square = 0.815 menyatakan besarnya koreksi terhadap R^2 sebesar 81.5 %.
- Std. Error of the Estimate = 1.71 merupakan ukuran besarnya variansi dalam model regresi adalah 1.71

2. **Uji Overall**

Syntax :

# Model Regresi
modelfix <-lm(Ytf3~`Rata-rata Suhu Ruangan`,data = data)

# Uji Overall
summary.aov(modelfix)

Output :

```
> # Uji Overall
> summary.aov(modelfix)
                         Df   Sum Sq  Mean Sq F value    Pr(>F)    
`Rata-rata Suhu Ruangan`  1 30416013 30416013   86.67 4.557e-10 ***
Residuals                28  9826376   350942                      
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>
```

- Hipotesis

  H_0 : semua β_i = 0, i = 0,1,2 (model regresi tidak layak digunakan)

  H_1 : ada $\beta_i \neq 0$, i = 0,1 (model regresi layak digunakan)

  dengan :

  β_0= parameter konstanta regresi

β_1 = parameter koefisien regresi

- Tingkat Signifikansi

  α = 0.05
- Statistik Uji

  P-Value = 4.557e-10
- Daerah Kritik

  H_0 ditolak jika P-Value $< \alpha$
- Kesimpulan

  Karena nilai P-Value = 4.557e-10 < 0,05 = α, maka H_0 ditolak .Sehingga dapat disimpulkan model regresi layak digunakan.
- Interpretasi

  Berdasarkan pengujian yang telah dilakukan dengan *software R.* dilakukan uji overall / keseluruhan dengan data jumlah cacat produksi selama 30 hari berturut-turut dan suhu ruangan, ingin diuji apakah model regresi layak digunakan atau tidak, dengan hipotesis H_0 : semua β_i = 0, i = 0,1,2 (model regresi tidak layak digunakan). Dengan tingkat signifikansi (α) sebesar 0.05 . Didapatkan hasil P-Value = 4.557e-10 < 0,05 = α, maka H_0 ditolak Sehingga dapat disimpulkan model regresi layak digunakan.

3. **Uji Parsial**

Syntax :

#Uji Parsial
summary(modelfix)

Output :

```
> #Uji Parsial
> summary(modelfix)

Call:
lm(formula = Ytf3 ~ `Rata-rata Suhu Ruangan`, data = data)

Residuals:
     Min       1Q   Median       3Q      Max 
-998.061 -423.848  -78.883  254.849 1374.868 

Coefficients:
                          Estimate Std. Error t value  Pr(>|t|)    
(Intercept)              -7864.787    985.815 -7.9780 1.091e-08 ***
`Rata-rata Suhu Ruangan`   392.071     42.114  9.3097 4.557e-10 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 592.4 on 28 degrees of freedom
Multiple R-squared:  0.75582,	Adjusted R-squared:  0.7471 
F-statistic:  86.67 on 1 and 28 DF,  p-value: 4.5566e-10
```

Uji Parsial Konstanta

- Hipotesis

  H_0 : β_0 = 0, (konstanta regresi tidak signifikan terhadap model regresi)

  H_1 : $\beta_0 \neq 0$, (konstanta regresi signifikan terhadap model regresi)

dengan :
β_0= parameter konstanta regresi

- Tingkat Signifikansi
  α = 0.05
- Statistik Uji
  P-Value = 1.091e-08
- Daerah Kritik
  H_0 ditolak jika P-Value $< \alpha$
- Kesimpulan
  Karena nilai P-Value = 1.091e-08 < 0,05 = α, maka H_0 ditolak. Sehingga dapat disimpulkan konstanta regresi signifikan terhadap model regresi.
- Interpretasi
  Berdasarkan pengujian yang telah dilakukan dengan *software R.* dilakukan Uji Parsial Konstanta dengan data jumlah cacat produksi selama 30 hari berturut-turut dan suhu ruangan, ingin diuji apakah konstanta regresi signifikan terhadap model regresi atau tidak, dengan hipotesis $H_0 : \beta_0 = 0$, (konstanta regresi tidak signifikan terhadap model regresi). Dengan tingkat signifikansi (α) sebesar 0.005 (nilai alpha). Didapatkan hasil P-Value = 1.091e-08 < 0,05 = α, maka H_0 ditolak. Sehingga dapat disimpulkan konstanta regresi signifikan terhadap model regresi.

```
> #Uji Parsial
> summary(modelfix)

Call:
lm(formula = Ytf3 ~ `Rata-rata Suhu Ruangan`, data = data)

Residuals:
     Min       1Q   Median       3Q      Max 
-998.061 -423.848  -78.883  254.849 1374.868 

Coefficients:
                          Estimate Std. Error t value  Pr(>|t|)    
(Intercept)              -7864.787    985.815 -7.9780 1.091e-08 ***
`Rata-rata Suhu Ruangan`   392.071     42.114  9.3097 4.557e-10 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 592.4 on 28 degrees of freedom
Multiple R-squared:  0.75582,	Adjusted R-squared:  0.7471 
F-statistic:  86.67 on 1 and 28 DF,  p-value: 4.5566e-10
```

Uji Parsial Koefisien

- Hipotesis
  $H_0 : \beta_1 = 0$, (Koefisien regresi tidak signifikan terhadap model regresi)
  $H_1 : \beta_1 \neq 0$, (Koefisien regresi signifikan terhadap model regresi)
  dengan :
  β_1= parameter koefisien variabel Rata-rata Suhu Ruangan
- Tingkat Signifikansi
  α = 0.05
- Statistik Uji

P-Value = 4.557e-10

- Daerah Kritik
  H_0 ditolak jika P-Value $< \alpha$
- Kesimpulan
  Karena nilai P-Value = 4.557e-10 < 0,05 = α, maka H_0 ditolak Sehingga dapat disimpulkan koefisien regresi variabel Rata-rata Suhu Ruangan signifikan terhadap model regresi.
- Interpretasi
  Berdasarkan pengujian yang telah dilakukan dengan *software R*. dilakukan Uji Parsial Koefisien dengan data jumlah cacat produksi selama 30 hari berturut-turut dan suhu ruangan, ingin diuji apakah koefisien variabel Rata-rata Suhu Ruangan signifikan terhadap model regresi atau tidak, dengan hipotesis H_0 : β_1 = 0, (Koefisien regresi tidak signifikan terhadap model regresi). Dengan tingkat signifikansi (α) sebesar 0.05. Didapatkan hasil P-Value = 4.557e-10 < 0,05 = α, maka H_0 ditolak Sehingga dapat disimpulkan koefisien regresi variabel Rata-rata Suhu Ruangan signifikan terhadap model regresi.

Interpretasi Keseluruhan :
(Gunakan bahasa dan pemahaman sendiri dalam menjelaskan semua uji yang telah dilakukan)

4. **Model Regresi**

Syntax :

```
# Model terbentuk
modelfix$coefficients
```

Output :

```
> # Model terbentuk
> modelfix$coefficients
            (Intercept) `Rata-rata Suhu Ruangan`
             -7864.7874                392.0711
```

Setelah dilakukan uji overall dan uji parsial dan didapatkan bahwa model layak digunakan seluruh konstanta dan koefisien variabel independen signifikan didapatkan model sebagai berikut:

$$y' = b_0 + b_1$$

Jumlah cacat produksi' = -7864.787 + 392.071*Rata - rata suhu ruangan

Interpretasi :

- setiap penambahan/ pengurangan 1 satuan variabel independen (rata - rata suhu ruangan) maka nilai variabel dependen (Jumlah cacat produksi) akan bertambah sebesar 392.071 satuan dengan asumsi variabel lain konstan.

- Jika variabel independen (rata - rata suhu ruangan) bernilai nol, maka nilai variabel dependen (Jumlah cacat produksi) adalah -7864.787 satuan (konstanta).

Karena kita telah melakukan transformasi terhadap y' dimana $y' = y^3$ maka jika dikembalikan ke bentuk awal diperoleh model regresi akhir sebagai berikut :

$$y^3 = b_0 + b_1$$

$$y = (b_0 + b_1)^{1/3}$$

Jumlah cacat produksi = (-7864.787 + 392.071*Rata - rata suhu ruangan)$^{1/3}$

G. Latihan Soal

1. Suatu perusahaan otomotif sedang memeriksa performa mobil baru produksinya, dimana sedang diperiksa jarak yang diperlukan mobil tersebut untuk berhenti (dalam feet atau ft) berdasarkan kecepatannya (diukur dalam mph atau miles per hour). Bantulah perusahaan tersebut membuat model yang tepat untuk memprediksi jarak diperlukan mobil untuk berhenti berdasarkan kecepatannya.
2. Badan klimatologi ingin melihat pengaruh suhu terhadap kecepatan angin, untuk itu dilakukan observasi terhadap suhu dan kecepatan angin di kota New York selama 153 hari berturut-turut. Lakukanlah analisis terhadap data yang telah didapatkan, kesimpulan apa yang diperoleh?

Note : untuk seluruh data telah tersedia dalam file excel dan gunakan tingkat signifikansi sebesar 5%