

- ▶ Statistika leverage

$$h_{ii} = \frac{1}{n} + \frac{(x_i - \bar{x})^2}{\sum_{k=1}^n (x_k - \bar{x})^2}, \quad i = 1, 2, \dots, n$$

- ▶ Standardized residual

$$R_i = \frac{E_i}{S \sqrt{1 - h_{ii}}}, \quad i = 1, 2, \dots, n.$$

- ▶ Deleted residual

$$\text{deleted residual } D_i = Y_i - \hat{Y}_{(i)}$$

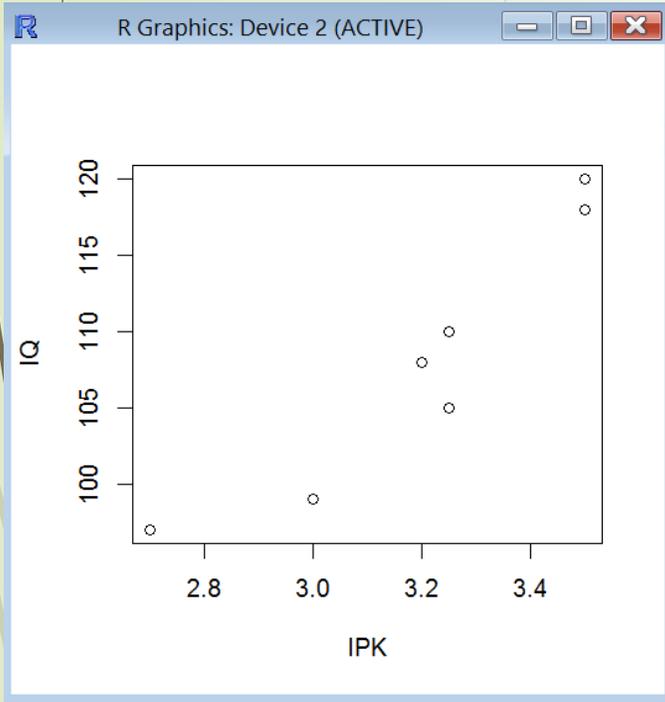
$$\text{Var}(D_i) = \frac{S_{(i)}^2}{1 - h_{ii}}, \quad i = 1, 2, \dots, n.$$

- ▶ studentized deleted residuals  $t_i$

$$t_i = \frac{D_i}{S_{(i)}/(1 - h_{ii})}, \quad i = 1, 2, \dots, n.$$

# Membuat plot

```
> u=data.frame(IPK, IQ)  
> plot(u,col=1)
```



```
> m1=lm(IPK~IQ,data=u)  
> summary(m1)
```

```
Call:  
lm(formula = IPK ~ IQ, data = u)  
  
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
```

# Mencari DFBETAS

- Untuk cek estimasi dan prediksi dari model
- Suatu observasi dianggap berpengaruh jika estimasi atau prediksinya terlalu besar/kecil
- Ukuran dalam estimasi parameter dengan DFBETAS. DFBETAS untuk intercept  $b_0$  dan  $b_1$  adalah :

$$(DFBETAS)_{0(i)} = \frac{b_0 - b_{0(i)}}{S_{(i)} \sqrt{\frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^n (x_i - \bar{x})^2}}}, \quad i = 1, 2, \dots, n.$$

$$(DFBETAS)_{1(i)} = \frac{b_1 - b_{1(i)}}{S_{(i)} \left[ \sum_{i=1}^n (x_i - \bar{x})^2 \right]^{-1/2}}, \quad i = 1, 2, \dots, n.$$

- Jika DFBETAS besar maka observasi mempunyai pengaruh besar terhadap koefisien regresi
- Observasi dianggap mencurigakan jika DFBETAS mempunyai besar lebih besar dari 1 untuk data kecil dan  $2/\sqrt{n}$  untuk data besar

```
> dfb<-dfbetas (fm)
> head(dfb)
      (Intercept)          IQ
1  0.464319910 -0.4366691319
2 -0.013842045  0.0147902931
3  0.001470352 -0.0002812224
4  0.003479708 -0.0051802139
5 -2.873052701  2.7324526463
6  0.488206346 -0.5159486020
```

# DFFITS

- Selain DFBETAS bisa juga diukur dari DFFITS

$$(DFFITS)_i = \frac{\hat{Y}_i - \hat{Y}_{(i)}}{S_{(i)} \sqrt{h_{ii}}}, \quad i = 1, 2, \dots, n.$$

- Observasi dicurigai jika nilai DFFIT lebih dari satu

```
> dff<-dffits(fm)
> dff[1:7]
      1          2          3          4          5          6
0.58327289  0.01913985  0.01594062 -0.02317005 -3.37758354 -0.62475614
      7
0.78778342
```

# Jarak Cook's

- Ukuran lain untuk mengukur data berpengaruh adalah jarak Cook's

$$D_i = \frac{E_i^2}{(p+1)S^2} \cdot \frac{h_{ii}}{(1-h_{ii})^2}, \quad i = 1, 2, \dots, n$$

- Hasil jarak Cook's dibandingkan dengan nilai quantile  $f(df1, df2=n-2)$  jika nilai  $D$  lebih besar dari  $Q2$  maka observasi dianggap ekstrim

```
> cooksD<-cooks.distance(fm)
```

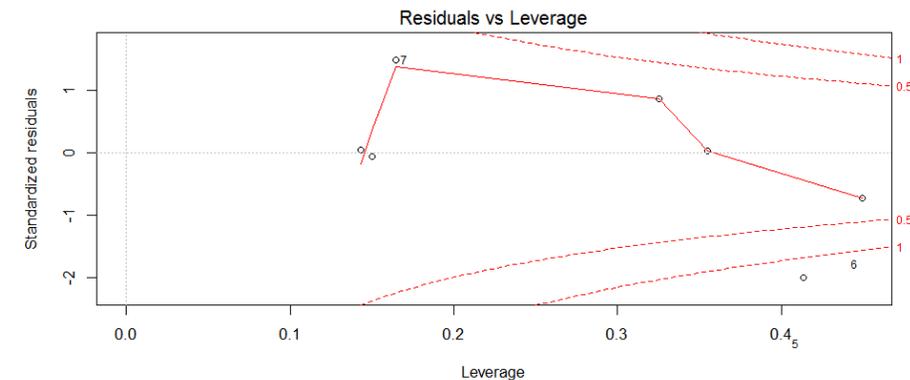
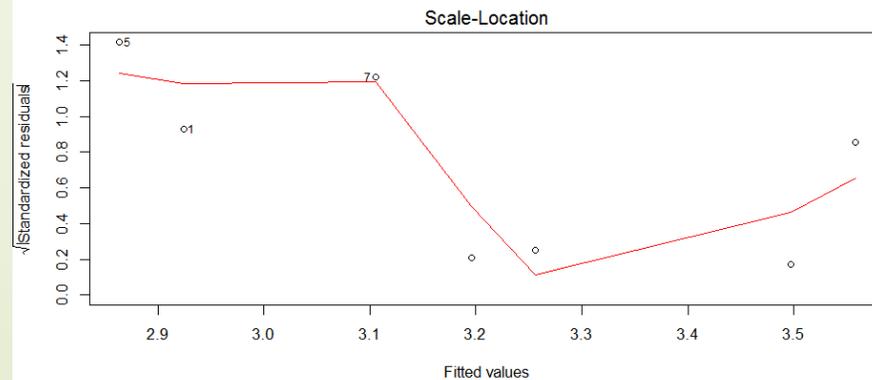
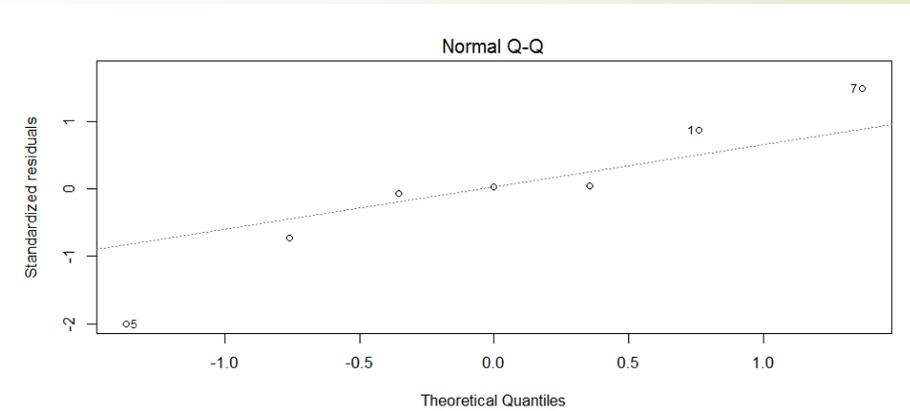
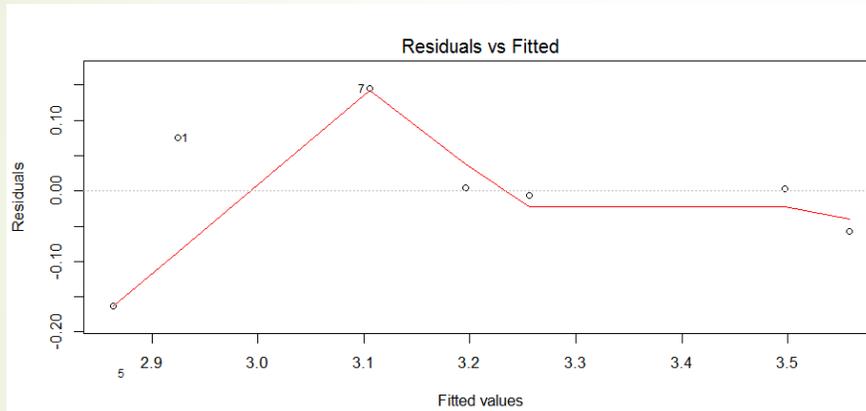
```
> cooksD[1:7]
```

```
      1          2          3          4          5          6  
0.1807120009 0.0002289205 0.0001587542 0.0003352777 1.4129852593 0.2178855503  
      7  
0.2168504879
```

# Identifikasi lain...

```
> par(mfrow=c(2,2))
> plot(fm)
```

```
> 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          7
0.86633904 0.02886657 0.04363896 -0.06155248 -2.00228192 -0.73090337
1.48480525
```



IQ (x)	IPK (y)
99	3.00
118	3.50
108	3.20
110	3.25
97	2.70
120	3.50
105	3.25

# Tugas Kelompok (TK) KD 2

## Analisis Korelasi & Regresi Berganda

- Presentasi minggu ke-2 dan ke-3, 2 kelompok/ minggu
- Review Teaching dan Material setiap setelah presentasi
- Paper maksimal 20 halaman
- Presentasi maksimal 30' / kelompok
- Semakin lengkap analisis semakin menambah poin TK
- Isi Paper:
  1. Pendahuluan
  2. Permasalahan
  3. Pembahasan
  4. Kesimpulan
  5. Lampiran



# Option Tema

1. Menganalisis korelasi dan regresi linier berganda
  - Uji asumsi pra analisis regresi
  - analisis korelasi
  - analisis regresi
  - Pemeriksaan sisa/ galat
2. Memilih model regresi linier berganda terbaik dengan metode:
  - Stepwise
  - Forward
  - Backward