

TUGAS KELOMPOK 15 ANREG

1. YUMNA SALMA (G1401221004) 2. CINDY APRILIYANTI (G1401221010) 3. Adinda Shabrina Putri Salsabila(G1401221081) 4. Ria Yunita(G1401221115)

2024-02-11

```
Data_kaggle <- read.csv("https://github.com/adindashabrina/DataAnreg/raw/main/jam_tidur.csv", header = TRUE, sep="," )
head(Data_kaggle)
```

##	Person.ID	Gender	Age	Occupation	Sleep.Duration	Quality.of.Sleep
## 1	1	Male	27	Software Engineer	6.1	6
## 2	2	Male	28	Doctor	6.2	6
## 3	3	Male	28	Doctor	6.2	6
## 4	4	Male	28	Sales Representative	5.9	4
## 5	5	Male	28	Sales Representative	5.9	4
## 6	6	Male	28	Software Engineer	5.9	4
##	Physical.Activity.Level	Stress.Level	BMI.Category	Blood.Pressure	Heart.Rate	
## 1	42	6	Overweight	126/83	77	
## 2	60	8	Normal	125/80	75	
## 3	60	8	Normal	125/80	75	
## 4	30	8	Obese	140/90	85	
## 5	30	8	Obese	140/90	85	
## 6	30	8	Obese	140/90	85	
##	Daily.Steps	Sleep.Disorder				
## 1	4200	None				
## 2	10000	None				
## 3	10000	None				
## 4	3000	Sleep Apnea				
## 5	3000	Sleep Apnea				
## 6	3000	Insomnia				

```
y<-Data_kaggle$Quality.of.Sleep
x<-Data_kaggle$Sleep.Duration
```

```
data_tidur <-data.frame(cbind(x,y))
head(data_tidur)
```

```
##      x y
## 1 6.1 6
## 2 6.2 6
## 3 6.2 6
## 4 5.9 4
## 5 5.9 4
## 6 5.9 4
```

```
n <- nrow(data_tidur)
n
```

```
## [1] 373
```

```
p<- ncol(data_tidur)
p
```

```
## [1] 2
```

PEMBENTUKAN MODEL TANPA FUNGSI BAWAAN(MANUAL)

Parameter Regresi

```
b1<- (sum(x*y) -sum(x)*sum(y)/n)/(sum(x^2) - (sum(x)^2/n))
b0<-mean(y) -b1*mean(x)
b1
```

```
## [1] 1.327017
```

```
b0
```

```
## [1] -2.152645
```

INTEPRETASI B1

$B1 = 1,32 > 0$ Koefisien 1.327 menunjukkan bahwa untuk setiap peningkatan satu jam dalam durasi tidur (x), kualitas tidur (y) meningkat sebesar 1.327. Selain itu, $B1$ ini juga menunjukkan hubungan yang positif antara durasi tidur dengan kualitas tidur.

INTERPRETASI B0

$B0 = -2,15 > 0$ Nilai $B0$ -2,15 menunjukkan bahwa jika durasi tidur nya adalah 0, maka kualitas tidur akan memiliki nilai -2,15. Namun, interpretasi ini tidak bermakna karena dalam konteks dunia nyata, durasi tidur tidak mungkin 0, sehingga saat $B0$ nya bernilai 0 tidak mengandung makna tertentu.

UKURAN KERAGAMAN

```
galat<-y-(b0+b1*x)
JKG <- sum((y - (b0+b1*x))^2)
JKG
```

```
## [1] 117.3708
```

```
JKReg <- sum(((b0+b1*x) - mean(y))^2)
JKReg
```

```
## [1] 414.1734
```

```
JKT <- sum((y - mean(y))^2)
JKT
```

```
## [1] 531.5442
```

```
dbReg<-1
dbg<-n-2
dbt<-n-1
dbReg
```

```
## [1] 1
```

```
dbg
```

```
## [1] 371
```

```
dbt
```

```
## [1] 372
```

```
Fhit<-(JKReg/dbReg)/(JKG/dbg)
Fhit
```

```
## [1] 1309.17
```

PERSAMAAN REGRESI LINEAR SEDERHANA

```
model<-lm(y~x,data<-data_tidur)
summary(model)
```

```
##
## Call:
## lm(formula = y ~ x, data = data <- data_tidur)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.67676 -0.19809  0.06732  0.40381  0.73082
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -2.15264    0.26310  -8.182 4.51e-15 ***
## x             1.32702    0.03668  36.182 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5625 on 371 degrees of freedom
## Multiple R-squared:  0.7792, Adjusted R-squared:  0.7786
## F-statistic: 1309 on 1 and 371 DF, p-value: < 2.2e-16
```

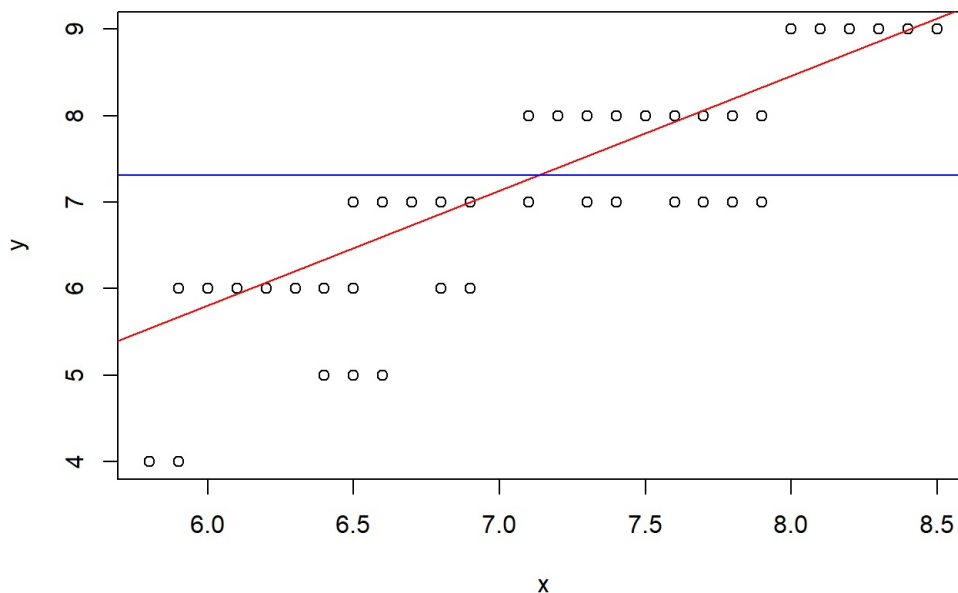
Berdasarkan perhitungan, didapat dugaan persamaan regresi linear sebagai berikut. Dugaan persamaan tersebut dapat pula disebut sebagai nilai harapan dari peubah respons Y y duga $= -2,15 + 1,327x$

INTERPRETASI PERSAMAAN REGRESI

persamaan regresi tersebut menunjukkan bahwa semakin lama seseorang tidur, kemungkinan besar kualitas tidurnya akan semakin baik atau kualitas tidurnya meningkat sebesar 1,327 per satu jam durasi tidur.

Scatter Plot

```
y.bar <- mean(y)
plot(x,y)
abline(model, col="red")
text(30, 220, "Y_duga", adj = c(-0.1, 1.5), col = "red", cex = 0.8)
abline(h=y.bar, col="blue")
text(31, 185, "Y_bar", adj = c(-0.1, 1.5), col = "blue", cex = 0.8)
```



INTERPRETASU SCATTER PLOT

garis biru menunjukkan rata-rata dari kualitas tidur, sedangkan garis merah merupakan garis persamaan linear sederhana yang bernilai positif karena garisnya yang menuju ke atas kanan dari sumbu y.

```
anova.model <- anova(model)
anova.model
```

```
## Analysis of Variance Table
##
## Response: y
##           Df Sum Sq Mean Sq F value    Pr(>F)
## x           1  414.17   414.17   1309.2 < 2.2e-16 ***
## Residuals  371  117.37     0.32
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
(KTG<-anova.model$'Mean Sq'[2])
```

```
## [1] 0.3163634
```

Dari tabel anova diatas diperoleh informasi bahwa nilai dari JKR nya sebesar 414,17, JKG sebesar 117,37, JKT sebesar 531,54, dan KTG sebesar 0,32, KTR 414,17.

Galat baku

```
galat_baku<- sqrt(KTG)
galat_baku
```

```
## [1] 0.5624619
```

KORELASI

```
r<- (sum(x*y)-sum(x)*sum(y)/n)/
sqrt((sum(x^2)-(sum(x)^2/n))*(sum(y^2)-(sum(y)^2/n)))
r
```

```
## [1] 0.8827168
```

INTERPRETASI

Korelasi 0.8827168 menunjukkan nilai kekuatan hubungan antara variabel penjelas yakni durasi tidur terhadap variabel respon yakni kualitas tidur yang baik karena mendekati nilai 1. Dari nilai ini juga dapat diketahui arah hubungan antara kedua peubah tersebut yakni linear positif, yang berarti semakin meningkatnya durasi tidur, maka kualitas tidur juga akan semakin baik.

Koefisien Determinasi

```
Koef_det<- r^2
Koef_det
```

```
## [1] 0.779189
```

INTERPRETASI

Koefisien determinasi 0.779189 menunjukkan kebaikan model regresi yang cukup baik dalam menggambarkan hubungan antara variabel penjelas dengan variabel respon.

ADJUSTED R SQUARE

```
Adj_R2<-1-((1-Koef_det)*(n-1)/(n-1-1))
Adj_R2
```

```
## [1] 0.7785938
```

INTERPRETASI

Adjusted R Square 0.7785938 atau sekitar 78% menunjukkan besar ukuran variabel penjelas dapat menjelaskan variabel respon, yang berarti sekitar 78% dari variasi dalam variabel respon dapat dijelaskan oleh variabel penjelas yang ada dalam model regresi.

UJI HIPOTESIS B0

```
(KTG<-anova.model$'Mean Sq'[2])
```

```
## [1] 0.3163634
```

```
(se_b0 <- sqrt(KTG*(1/n+mean(x)^2/sum((x-mean(x))^2))))
```

```
## [1] 0.263096
```

```
(t_b0 <- b0/se_b0)
```

```
## [1] -8.181975
```

PENJELASAN UJI HIPOTESIS B0

Uji hipotesis untuk intersep garis regresi (beta0) H0: beta0 = 0 (semua nilai y dapat dijelaskan oleh x) H1: beta0 != 0 (ada nilai y yang tidak dapat dijelaskan oleh x)

t-hitung = -8.182, t-tabel = 1,966378803 (t(alfa/2,db =n-2)) Karena t-hitung < -t-tabel maka tolak H0 yang artinya terdapat cukup bukti untuk menyatakan semua nilai kualitas tidur dapat dijelaskan oleh durasi total waktu tidur pada selang kepercayaan 95%.

UJI HIPOTESIS B1

```
(se_b1 <- sqrt(KTG/sum((x-mean(x))^2)))
```

```
## [1] 0.03667571
```

```
(t_b1 <- b1/se_b1)
```

```
## [1] 36.18245
```

PENJELASAN UJI HIPOTESIS B1

Uji hipotesis kemiringan garis regresi populasi (beta1) H0: beta1 = 0 (Tidak ada hubungan linier antara kualitas tidur dengan lama waktu tidur) H1: beta1 != 0 (Ada hubungan linier antara kualitas tidur dengan lama waktu tidur)

t-hitung = 36.18245, t-tabel = 1,966378803 (t(alfa/2,db =n-2)) Karena t-hitung > t-tabel maka Tolak H0 yang artinya terdapat cukup bukti untuk menyatakan adanya hubungan linier antara kualitas tidur dengan lama waktu tidur pada selang kepercayaan 95%.

PENDUGA SELANG KEPERCAYAAN 95% BAGI B0

```
#batas bawah beta0  
(ba.b0 <- b0 - abs(qt(0.025, df=n-2))*se_b0)
```

```
## [1] -2.669991
```

```
#batas atas beta0  
(ba.b0 <- b0 + abs(qt(0.025, df=n-2))*se_b0)
```

```
## [1] -1.635298
```

Pada tingkat kepercayaan 95% dapat dinyatakan dugaan parameter beta0 berada dalam selang -2.67 hingga -1.635.

PENDUGA SELANG KEPERCAYAAN 95% BAGI B1

```
#batas bawah beta1  
(bb.b1 <- b1 - abs(qt(0.025, df=n-2))*se_b1)
```

```
## [1] 1.254899
```

```
#batas atas beta1  
(ba.b1 <- b1 + abs(qt(0.025, df=n-2))*se_b1)
```

```
## [1] 1.399135
```

Pada tingkat kepercayaan 95% dapat dinyatakan dugaan parameter β_1 berada dalam selang 1.2549 hingga 1.399135.

SELANG KEPERCAYAAN RATAAN(NILAI HARAPAN) AMATAN

Misalkan kita ingin menduga nilai rataan (harapan) amatan ketika nilai $x=6.1$. Kita dapat memanfaatkan fungsi `predict` untuk mendapatkan nilai dugaan dan selang kepercayaannya sebagai berikut.

```
amatan.diduga <- data.frame(x=6.1)
predict(model, amatan.diduga, interval = "confidence")
```

```
##           fit          lwr          upr
## 1 5.942159 5.848394 6.035924
```

Berdasarkan output sintaks di atas, diketahui bahwa dugaan rataan nilai Y ketika nilai $x=6.1$ adalah 5.942159. Selain itu, diindikasikan juga bahwa dalam taraf kepercayaan 95%, diyakini bahwa nilai dugaan rataan Y ketika nilai $x=6.1$ berada dalam selang 5.848394 hingga 6.035924.

SELANG KEPERCAYAAN INDIVIDU AMATAN

Misalkan kita ingin menduga nilai individu amatan ketika nilai $x=6.1$. Kita dapat memanfaatkan fungsi `predict` untuk mendapatkan nilai dugaan dan selang kepercayaannya sebagai berikut.

```
predict(model, amatan.diduga, interval = "prediction")
```

```
##           fit          lwr          upr
## 1 5.942159 4.832179 7.05214
```

Berdasarkan output sintaks di atas, diketahui bahwa dugaan nilai individu Y ketika nilai $x=6.1$ adalah 5.942159. Selain itu, diindikasikan juga bahwa dalam taraf kepercayaan 95%, diyakini bahwa nilai amatan individu Y ketika nilai $x=6.1$ berada dalam selang 4.832179 hingga 7.05214.