

# PML\_M10\_16231059

Zafyra Nur Rizqi

2025-04-23

## 1. Uji Heteroskedastisitas

---

- Model Regresi

```
model <- lm(sales ~ youtube, data = marketing)
```

- Manual

```
resid_kuadrat <- residuals(model)^2
mod_resid <- lm(resid_kuadrat ~ marketing$youtube)
chi_sq <- nrow(marketing) * summary(mod_resid)$r.squared
p_val <- 1 - pchisq(chi_sq, df = 1)
cat("BP =", round(chi_sq, 3), "; p-value =", format.pval(p_val, digits = 4))
```

```
## BP = 48.038 ; p-value = 4.18e-12
```

- Fungsi langsung di R

```
bptest(model)
```

```
##
## studentized Breusch-Pagan test
##
## data: model
## BP = 48.038, df = 1, p-value = 4.18e-12
```

- Kesimpulan

Pendekatan manual melalui regresi kuadrat residual terhadap variabel independen serta metode statistik menggunakan fungsi `bptest()` dari paket `lmtest` menghasilkan nilai statistik Breusch-Pagan (BP) sebesar 48.038 dengan p-value yang sangat kecil (4.18e-12). Kesamaan hasil dari kedua metode tersebut menunjukkan adanya indikasi kuat terhadap heteroskedastisitas dalam model regresi. Oleh karena itu, asumsi mengenai homogenitas varians residual tidak terpenuhi pada model ini.

## 2. Uji Normalitas

---

- Model Regresi

```
log.model <- lm(log(sales) ~ youtube, data = marketing)
```

- Manual

```
error <- residuals(log.model)
e_cdf <- ecdf(error)
t_cdf <- pnorm(error, mean = 0, sd = sd(error))
D <- max(abs(e_cdf(error) - t_cdf))
print(D)
```

```
## [1] 0.0658037
```

- Fungsi langsung di R

```
ks.test(error, "pnorm", 0, sd(error))
```

```
##
## Asymptotic one-sample Kolmogorov-Smirnov test
##
## data: error
## D = 0.070804, p-value = 0.2686
## alternative hypothesis: two-sided
```

```
ks.test(error, "pnorm", 0, sd(residuals(model)))
```

```
##
## Asymptotic one-sample Kolmogorov-Smirnov test
##
## data: error
## D = 0.45656, p-value < 2.2e-16
## alternative hypothesis: two-sided
```

- Kesimpulan

Perhitungan manual menunjukkan nilai K-S sebesar 0,0658037, sementara hasil uji menggunakan fungsi `ks.test()` pada model logaritmik (`log.model`) menghasilkan nilai K-S sebesar 0,070804 dengan p-value 0,2686. Kedua nilai ini cukup serupa dan menunjukkan bahwa tidak ada perbedaan signifikan antara distribusi residual dan distribusi normal, sehingga asumsi normalitas dapat diterima pada model ini. Sebaliknya, uji K-S pada residual model tanpa transformasi (`model`) menghasilkan nilai statistik 0,45656 dan p-value yang sangat kecil ( $< 2.2e-16$ ), menandakan adanya penyimpangan signifikan dari distribusi normal. Dari perbedaan hasil ini dapat disimpulkan bahwa dengan penggunaan transformasi logaritmik dalam `log.model`, yang berfungsi untuk menstabilkan varians dan memperbaiki kemencengan data penjualan, sehingga menghasilkan residual yang lebih mendekati distribusi normal dan memenuhi asumsi dasar regresi.

### 3. Uji Autokorelasi

---

- Manual

```
n <- dim(marketing)[1]
r <- cor(error[-1], error[-n])
DW <- 2*(1-r)
DW
```

```
## [1] 1.819105
```

```
dU <- 1.778 # Memeriksa apakah nilai DW berada dalam kisaran yang menunjukkan tidak ada autokorelasi
if(DW > dU && DW < 4-dU){
  cat("No Autocorrelation, DW = ", DW)
} else {
  cat("Autocorrelation exists, DW = ", DW)
}
```

```
## No Autocorrelation, DW = 1.819105
```

- Fungsi langsung di R

```
dwtest(model) # Model linear biasa
```

```
##
## Durbin-Watson test
##
## data: model
## DW = 1.9347, p-value = 0.3213
## alternative hypothesis: true autocorrelation is greater than 0
```

```
dwtest(log.model) # Model logaritmik
```

```
##
## Durbin-Watson test
##
## data: log.model
## DW = 1.81, p-value = 0.08847
## alternative hypothesis: true autocorrelation is greater than 0
```

- Kesimpulan

Hasil uji autokorelasi baik dengan pendekatan manual (Durbin-Watson) maupun fungsi `dwtest()` menghasilkan kesimpulan yang sejalan. Pada model dengan transformasi logaritmik, nilai Durbin-Watson yang diperoleh adalah 1,819 dengan p-value 0,08847, sementara pada model tanpa transformasi, nilai Durbin-Watson mencapai 1,9347 dengan p-value 0,3213. Kedua nilai ini terletak di antara batas atas dan bawah yang kritis, yang berarti tidak ada cukup bukti untuk menyimpulkan adanya autokorelasi dalam residual. Ini menunjukkan bahwa asumsi independensi residual dapat diterima pada kedua model tersebut.

## 4. Uji Multikolinearitas (VIF)

---

- Manual

```
# Membuat model regresi untuk setiap variabel sebagai dependen
m_youtube <- lm(youtube ~ facebook + newspaper, data = marketing)
m_facebook <- lm facebook ~ youtube + newspaper, data = marketing)
m_newspaper <- lm(newspaper ~ youtube + facebook, data = marketing)

# Mengambil R-squared (bukan adjusted R-squared) dari setiap model
R2_youtube <- summary(m_youtube)$r.squared # Menggunakan R-squared
R2_facebook <- summary(m_facebook)$r.squared # Menggunakan R-squared
R2_newspaper <- summary(m_newspaper)$r.squared # Menggunakan R-squared

# Menghitung VIF untuk setiap model
VIF_y <- 1 / (1 - R2_youtube)
VIF_f <- 1 / (1 - R2_facebook)
VIF_n <- 1 / (1 - R2_newspaper)

# Menampilkan hasil dalam data frame
data.frame(Youtube = VIF_y,
           Facebook = VIF_f,
           Newspaper = VIF_n)
```

```
##      Youtube Facebook Newspaper
## 1 1.004611 1.144952 1.145187
```

- Fungsi langsung di R

```
model.mlr <- lm(sales ~ youtube + facebook + newspaper, data = marketing)
vif(model.mlr)
```

```
##      youtube  facebook newspaper
## 1.004611 1.144952 1.145187
```

- Kesimpulan

Perhitungan VIF secara manual dan menggunakan fungsi `vif()` dari paket `car` menghasilkan hasil yang konsisten. Semua nilai VIF berada dalam rentang 1,00–1,15, yang jauh di bawah ambang batas toleransi umum sebesar 10. Ini menunjukkan bahwa tidak ada masalah multikolinearitas antar variabel independen dalam model regresi berganda. Dengan demikian, model ini memenuhi asumsi bahwa tidak ada korelasi tinggi antar prediktor.