

Nama : Hasna Septinna  
NPM : 2113031033

1) Asumsi Linearitas

⇒ Manfaatnya untuk menjamin bahwa hubungan antar variabel dependen dan independen bersifat linear, sehingga model regresi menghasilkan data yang benar

⇒ Cara mengujinya :

⇒ Plot scatterplot antar variabel dependen dengan masing-masing variabel independen

⇒ Gunakan Uji Ramsey RESET (Regression Specification Error Test)

2) Asumsi Non Multikolinearitas

⇒ Manfaatnya untuk menghindari kesulitan dalam menginterpretasikan koefisien regresi dan meningkatkan stabilitas estimasi.

⇒ Cara Menguji :

⇒ Menghitung VIF, jika  $VIF > 10$ , maka multikolinearitas

⇒ Uji Korelasi antar variabel ~~in~~ independen.

3) Asumsi Homoskedastisitas

⇒ Manfaatnya untuk varians residual yang konstan memastikan bahwa efisiensi estimasi koefisien terjamin

⇒ Cara Menguji :

⇒ Plot residual vs nilai prediksi untuk melihat pola Heteroskedastisitas

⇒ Uji Breusch-Pagan atau Uji White

4) Asumsi Non Autokorelasi

⇒ Manfaatnya untuk menghindari masalah dimana residual dari model bertetapan satu sama lain, menyebabkan estimasi tidak efisien

⇒ Cara Menguji :

⇒ Uji Durbin Watson

⇒ Uji Breusch-Godfrey

5) Asumsi Normalitas Residual

⇒ Manfaatnya diperlukan untuk pengujian hipotesis dan pembuatan interval kepercayaan akurat.

⇒ Cara Menguji :

⇒ Plot histogram atau Q-Q plot dari residual

⇒ Uji Jarque-Bera atau Uji Shapiro Wilk

6) Asumsi non variabel yang terlewat :

⇒ Manfaatnya memastikan bahwa semua variabel relevan dimasukkan ke dalam model sehingga tidak ada bias spesifikasi

⇒ Cara Menguji :

⇒ Uji Ramsey RESET untuk mendeteksi spesifikasi model yang salah

7) Asumsi Independen dan Identik Terdistribusi (i.i.d) Data :



⇒ Mengingatnya memastikan bahwa data berasal dari populasi yang sama dan estimasi yang valid

⇒ Cara Menguji :

⇒ Dapat dilakukan dengan memahami konteks pengambilan sampel dan desain penelitian.

## 2. Penyebab Ketidakvalidan data :

1.) Kesalahan Pengukuran → Alat Ukur/Metode pengumpulan data yang tidak akurat  
→ Kesalahan input data.

2) Data Hilang → Responden tidak menjawab semua pertanyaan survey  
→ Data tidak tercatat baik dalam sistem

3) Outlier → Nilai yang berbeda yang dapat terjadi karena kesalahan input atau ekstrem

4) Kesalahan Pengkodean → Salah memasukkan kode kategori yang tidak konsisten

5) Duplikasi Data → Data tercatat lebih dari satu kali dalam dataset

Metode Mendeteksi :

1). Pemeriksaan Deskriptif → Mean, Median, Modus

2) Pemeriksaan Outlier → Boxplot atau scatterplot  
→ Z-Score atau IQR

3) Analisis Missing Data → Tabel frekuensi, Metode Visual

4) Validasi Cross-Referensi → Memeriksa data terhadap sumber data lain.

5) Pemeriksaan Konsistensi → Menguji Konsistensi Logis antar variabel

Penyebab Ketidaknormalan Data

1). Distribusi Asimetris → Data tidak berdistribusi normal

2). Outlier → Nilai ekstrem yang mengganggu distribusi normal

3) Multimodalitas → Data lebih dari satu puncak dalam distribusinya

4) Keterbatasan Sampel

Metode Mendeteksi Ketidaknormalan Data

1) Plot Histogram

2) Q-Q Plot

3) Uji Normalitas

4) Skewness dan Kurtosis

## 3. Regresi Linear Sederhana adalah analisis data statistik untuk mengidentifikasi hubungan antara satu variabel dependen (Y) dan satu variabel independen (X)

Regresi Linear Berganda adalah perluasan dari regresi linear sederhana, dimana hubungan antara satu variabel dependen (Y) dan dua atau lebih variabel independen ( $X_1, X_2, \dots$ )

Tahap Mengolah Data Regresi :

1.) Pengumpulan Data

2.) Pembersihan Data



- 3) Eksplorasi Data
- 4) Pembagian Data
- 5) Pemilihan Model
- 6) Estimasi Model
- 7) Uji Asumsi Klasik