

UJIAN AKHIR SEMESTER
MATA KULIAH : STATISTIK DAN PROBABILITAS
PROGRAM STUDI SISTEM INFORMASI

Nama : I Made Aditya Pradnyana Putra

NIM : 2415091024

Kelas : 1 DPS

Analisis Regresi Linear Sederhana Curah Hujan dan Hari Hujan di Provinsi Jawa Barat Tahun 2017 dengan menggunakan Program R

Error! Hyperlink reference not valid.

1. Data (*Jumlah Curah Hujan dan Hari Hujan di Provinsi Jawa Barat tahun 2017*)

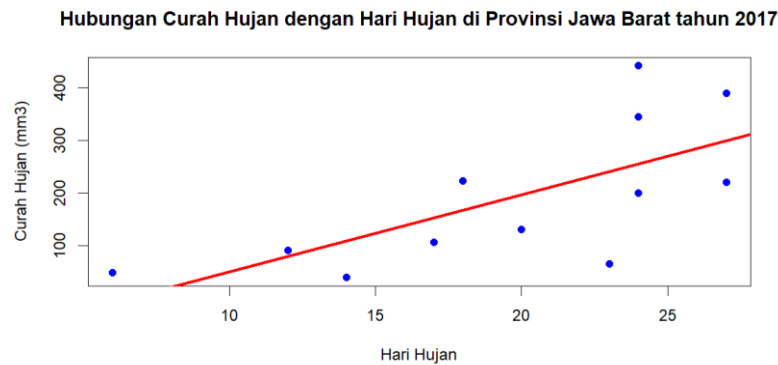
Bulan	Curah Hujan (m ³)	Hari Hujan
Januari	65	23
Februari	199,3	24
Maret	389,3	27
April	220	27
Mei	222,3	18
Juni	106,4	17
Juli	39,1	14
Agustus	48	6
September	91	12
Oktober	345	24
November	442	24
Desember	129,9	20

Data tersebut berisi informasi mengenai jumlah curah hujan yang tercatat setiap bulan berdasarkan jumlah hari hujan dalam kurun waktu satu tahun. Setiap bulan, jumlah hari hujan dapat mempengaruhi total curah hujan yang tercatat, dan analisis lebih lanjut dapat digunakan untuk memprediksi curah hujan berdasarkan hari hujan.

2. Analisis Regresi Linear Sederhana dengan Uji Asumsi

2.1. Uji Linearitas

Uji Linearitas dilakukan dengan menggunakan fungsi plot untuk mengetahui apakah hubungan antara variabel independen (Hari Hujan) dan variabel dependen (Curah Hujan) bersifat linear atau tidak.



Pada program R tersebut, diketahui bahwa hubungan antara variabel independen (Hari Hujan) dan variabel dependen (Curah Hujan) bersifat linear, karena titik-titik data tersebut mengikuti pola garis lurus (berwarna merah) yang terdapat pada plot (secara langsung), maka hubungan tersebut dianggap linear.

Tahapan-tahapan Uji Linearitas:

- Membuat sebuah plot data tentang Hubungan Curah Hujan dengan Hari Hujan di Provinsi Jawa Barat tahun 2017 yang berisikan data Hari Hujan (x) dan Curah Hujan (y) dengan menggunakan fungsi `plot()`, untuk memberikan gambaran visual agar hubungan antara kedua variabel terlihat linear.
- Menambahkan garis regresi pada plot menggunakan fungsi `abline()`, untuk menunjukkan garis terbaik yang menggambarkan hubungan linear antara Hari Hujan dan Curah Hujan, serta menunjukkan prediksi jumlah Curah Hujan untuk setiap jumlah Hari Hujan sesuai dengan model regresi yang dibangun.

2.2. Uji Normalitas

Pengujian ini bertujuan untuk memastikan bahwa sisa perhitungan dari model regresi (residual) terdistribusi secara normal. Artinya, memastikan bahwa tidak ada pola yang ambigu dalam kesalahan perhitungan yang dapat mempengaruhi hasil dari model.

Tahapan-tahapan Uji Normalitas:

- Membangun Model Regresi Linear

Pada program R, model regresi sederhana dibangun untuk mengevaluasi hubungan antara variabel independen (Hari Hujan) dan variabel dependen (Curah Hujan). Dalam fungsinya, model ini menghasilkan residual, yaitu selisih antara nilai aktual dan nilai prediksi. Residual digunakan untuk memeriksa apakah asumsi regresi, termasuk normalitas residual, terpenuhi.

b. Membangun Histogram Residual

Dalam program R yang telah dibuat, terdapat tampilan distribusi residual secara visual berupa Histogram Residual yang menampilkan selisih antara nilai yang diprediksi dan nilai aktual dari model regresi. Histogram tersebut berbentuk simetris menyerupai distribusi normal (membentuk kurva lonceng), yang berarti asumsi normalitas dapat terpenuhi.

c. Q-Q Plot Residual

Pada program R terdapat visual dari Q-Q Plot Residual yang mengevaluasi normalitas residual dengan membandingkan distribusi residual terhadap distribusi normal. Pada titik-titik yang berada pada Q-Q plot, titik-titik tersebut mengikuti garis lurus (garis merah), yang berarti residual dapat dianggap terdistribusi normal.

d. Uji Normalitas Residual dengan Shapiro-Wilk

Uji ini dilakukan untuk menentukan apakah residual terdistribusi normal atau tidak. Pada program R, hasil uji yang di dapat menunjukkan bahwa $p\text{-value} = 0.787$, yang berarti bahwa $p\text{-value}$ (nilai p) > 0.05 , sehingga tidak dapat menolak hipotesis nol, dengan kata lain residual terdistribusi normal.

2.3. Homogenitas

Uji homogenitas digunakan untuk memastikan bahwa variasi atau penyebaran kesalahan perhitungan (residual) tetap konstan di seluruh data. Hasilnya menunjukkan bahwa penyebaran kesalahan perhitungan cukup seragam, yang berarti tidak ada masalah dalam pemodelan terkait variasi kesalahan di berbagai titik data.

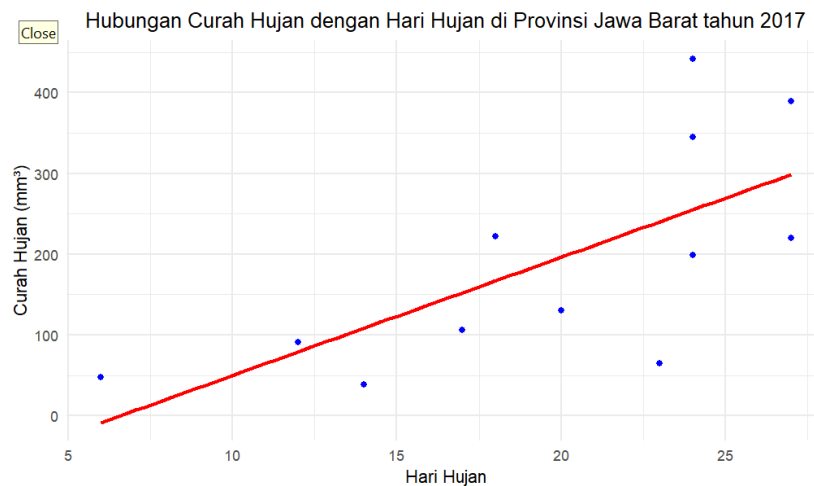
a. Plot Residual vs Fitted

Plot ini digunakan untuk memeriksa apakah variansi residual bersifat homogen (atau seragam) pada seluruh rentang nilai prediksi (fitted values). Pada program R, plot Residual vs Fitted yang di tampilkan bersifat homogen (seragam) pada seluruh rentang nilai prediksinya

b. Uji Breusch-Pagan

Uji ini digunakan untuk memeriksa apakah terdapat heteroskedastisitas (ketidakhomogenan variansi residual). Jika $p\text{-value} > 0.05$, maka asumsi homogenitas variansi terpenuhi. Pada program R, hasil uji asumsi yang di dapat adalah $p\text{-value} = 0.1633$, itu berarti $p\text{-value} > 0.05$, maka asumsi homogenitas variansi dapat terpenuhi

3. Visualisasi



Dalam visualisasi yang dibuat menggunakan ggplot, grafik memberikan gambaran yang jelas tentang hubungan antara curah hujan dan hari hujan. Titik-titik biru mewakili data (jumlah hari hujan dan curah hujan) yang tersebar di sekitar garis merah, yang menunjukkan kecenderungan linear. Garis regresi merah tersebut cukup baik dalam menggambarkan hubungan antara kedua variabel, menunjukkan tren bahwa semakin banyak hari hujan, maka semakin tinggi curah hujan yang tercatat. Hal ini menandakan bahwa hubungan antara jumlah hari hujan dan curah hujan di Provinsi Jawa Barat cenderung linear, sehingga asumsi linearitas terpenuhi. Visualisasi ini secara jelas menunjukkan adanya kecenderungan positif antara jumlah hari hujan dan curah hujan, yang mendukung hasil analisis regresi serta uji asumsi linearitas.

4. Interpretasi

Pada analisis regresi yang telah dilakukan terhadap data jumlah curah hujan dan hari hujan di Provinsi Jawa Barat tahun 2017, terdapat hubungan yang jelas antara kedua variabel, baik variabel independen (Hari Hujan) dan variabel dependen (Curah Hujan). Hasil analisis menunjukkan bahwa semakin banyak hari hujan dalam sebulan, maka semakin tinggi jumlah curah hujan yang tercatat. Uji linearitas mengonfirmasi bahwa hubungan antara jumlah hari hujan dan curah hujan bersifat linear. Selain itu, uji normalitas terhadap residual menunjukkan bahwa kesalahan model terdistribusi secara normal, yang menerangkan bahwa model yang dibangun valid. Uji homogenitas juga memperlihatkan bahwa variasi kesalahan perhitungan tetap konsisten di seluruh data, yang berarti model regresi dapat diandalkan. Secara keseluruhan, hasil ini memberikan gambaran yang jelas bahwa jumlah hari hujan berperan penting dalam menentukan curah hujan.