

ABSTRAK

Nama : Fadila Rahmawati
NIM : 1207010018
Judul : Metode Deteksi *Outlier* Menggunakan Estimasi *Minimum Covariance Determinant* dan *Least Trimmed Squares* untuk Data Pendidikan di Jawa Barat

Outlier merupakan nilai data yang menyimpang jauh dari pola umum dan dapat merusak keakuratan model regresi. Metode Kuadrat Terkecil (MKT) sangat *sensitive* terhadap *Outlier* sehingga dapat menghasilkan estimasi parameter yang bias. Untuk mengatasi hal ini, digunakan pendekatan regresi *Robust* yang lebih tahan terhadap *Outlier*. Penelitian ini bertujuan untuk mendeteksi dan menangani outlier pada data Pendidikan di Jawa Barat tahun 2024 dengan menggunakan metode *Minimum Covariance Determinant* (MCD) dan *Least Trimmed Squares* (LTS) yang diterapkan pada data pendidikan di Jawa Barat untuk mengidentifikasi outlier pada kabupaten atau kota. Data yang digunakan meliputi 27 kabupaten/kota dengan variabel : Indeks Pendidikan sebagai variabel dependen, serta PDRB per kapita, tingkat kemiskinan, jumlah sekolah, jumlah guru, kepadatan penduduk sebagai variabel independen. Proses analisis dilakukan melalui dua tahap, yaitu pendeteksian outlier menggunakan metode MCD, kemudian estimasi model regresi dengan metode MKT setelah data *outlier* dihapus, serta menggunakan metode LTS sebagai pembanding. Hasil penelitian menunjukkan bahwa metode MCD berhasil mengidentifikasi 10 *outlier* dan menghasilkan nilai *Mean Square Error* (MSE) sebesar 0,0005 lebih kecil dibandingkan metode LTS yang memiliki nilai sebesar 0,001. Dengan demikian, penggunaan metode MCD untuk mendeteksi dan menghapus outlier, diikuti dengan estimasi regresi linear berganda menggunakan MKT, terbukti meningkatkan akurasi dan kestabilan model regresi.

Kata Kunci : *Outlier*, Regresi *Robust*, *Minimum Covariance Determinant*, *Least Trimmed Squares*, *Mean Square Error*, Pendidikan Jawa Barat.

ABSTRACT

Nama : Fadila Rahmawati
NIM : 1207010018
Judul : Metode Deteksi Outlier Menggunakan Estimasi Minimum Covariance Determinant dan *Least Trimmed Squares* untuk Data Pendidikan di Jawa Barat

Outliers are data points that deviate significantly from the general pattern and can adversely affect the accuracy of regression models. The Ordinary Least Squares (OLS) method is highly sensitive to outliers, which can lead to biased parameter estimates. To address this issue, robust regression approaches that are more resistant to outliers are employed. This study aims to detect and handle outliers in educational data from West Java in 2024 using the Minimum Covariance Determinant (MCD) and Least Trimmed Squares (LTS) methods to identify outliers among districts or cities. The data comprises 27 districts/cities, with the Education Index as the dependent variable, and per capita Gross Regional Domestic Product (GRDP), poverty rate, number of schools, number of teachers, and population density as independent variables. The analysis is conducted in two stages: first, outlier detection using the MCD method; second, regression model estimation using OLS after removing outliers detected by MCD, with LTS employed as a comparative robust regression method. The results show that the MCD method successfully identified 10 outliers and achieved a Mean Square Error (MSE) of 0.0005, which is lower than the MSE of 0.001 produced by the LTS method. Therefore, using MCD to detect and remove outliers followed by multiple linear regression estimation with OLS improves the accuracy and stability of the regression model.

Keywords: Outlier, Robust Regression, Minimum Covariance Determinant, Least Trimmed Squares, Mean Square Error, West Java Education.