1. **MENGATASI MISSING VALUE**

OLS Regression Results

==============================================================================

Dep. Variable: y R-squared: 0.195

Model: OLS Adj. R-squared: 0.195

Method: Least Squares F-statistic: 3549.

Date: Sat, 09 Mar 2019 Prob (F-statistic): 0.00

Time: 15:49:30 Log-Likelihood: -2.5454e+05

No. Observations: 43824 AIC: 5.091e+05

Df Residuals: 43820 BIC: 5.091e+05

Df Model: 3

Covariance Type: nonrobust

==============================================================================

coef std err t P>|t| [0.025 0.975]

------------------------------------------------------------------------------

Intercept 1609.8395 71.541 22.502 0.000 1469.618 1750.061

x[0] 4.3651 0.049 88.198 0.000 4.268 4.462

x[1] -5.8985 0.065 -90.272 0.000 -6.027 -5.770

x[2] -1.4223 0.070 -20.338 0.000 -1.559 -1.285

==============================================================================

Omnibus: 16279.775 Durbin-Watson: 0.112

Prob(Omnibus): 0.000 Jarque-Bera (JB): 82111.750

Skew: 1.738 Prob(JB): 0.00

Kurtosis: 8.735 Cond. No. 1.89e+05

==============================================================================

Dari output diatas diketahui untuk model 1 dengan hanya mengatasi missing value data. Diperoleh nilai R-squared sebesar 0,195

1. **Mengatasi Missing Value dan Outlier**

OLS Regression Results

==============================================================================

Dep. Variable: y2 R-squared: 0.206

Model: OLS Adj. R-squared: 0.206

Method: Least Squares F-statistic: 3544.

Date: Sat, 09 Mar 2019 Prob (F-statistic): 0.00

Time: 15:56:34 Log-Likelihood: -2.3118e+05

No. Observations: 40995 AIC: 4.624e+05

Df Residuals: 40991 BIC: 4.624e+05

Df Model: 3

Covariance Type: nonrobust

==============================================================================

coef std err t P>|t| [0.025 0.975]

------------------------------------------------------------------------------

Intercept 1274.1673 62.317 20.446 0.000 1152.024 1396.311

x2[0] 3.9602 0.043 91.040 0.000 3.875 4.045

x2[1] -4.9743 0.057 -86.619 0.000 -5.087 -4.862

x2[2] -1.1065 0.061 -18.163 0.000 -1.226 -0.987

==============================================================================

Omnibus: 6332.990 Durbin-Watson: 0.124

Prob(Omnibus): 0.000 Jarque-Bera (JB): 10159.534

Skew: 1.066 Prob(JB): 0.00

Kurtosis: 4.183 Cond. No. 1.88e+05

==============================================================================

Dari output diatas jika dilakukan pengujian untuk mengetahui outlier dan outlier diatasi diperoleh mode seperti diatas dengan R-squared sebesar 0,206.

1. **Mengatasi Missing Value dan transformasi**

OLS Regression Results

==============================================================================

Dep. Variable: y3 R-squared: 0.195

Model: OLS Adj. R-squared: 0.195

Method: Least Squares F-statistic: 3549.

Date: Sat, 09 Mar 2019 Prob (F-statistic): 0.00

Time: 16:06:46 Log-Likelihood: -2.5454e+05

No. Observations: 43824 AIC: 5.091e+05

Df Residuals: 43820 BIC: 5.091e+05

Df Model: 3

Covariance Type: nonrobust

==============================================================================

coef std err t P>|t| [0.025 0.975]

------------------------------------------------------------------------------

Intercept 1609.8395 71.541 22.502 0.000 1469.618 1750.061

x3[0] 4.3651 0.049 88.198 0.000 4.268 4.462

x3[1] -5.8985 0.065 -90.272 0.000 -6.027 -5.770

x3[2] -1.4223 0.070 -20.338 0.000 -1.559 -1.285

==============================================================================

Omnibus: 16279.775 Durbin-Watson: 0.112

Prob(Omnibus): 0.000 Jarque-Bera (JB): 82111.750

Skew: 1.738 Prob(JB): 0.00

Kurtosis: 8.735 Cond. No. 1.89e+05

==============================================================================

Dari output diatas jika dilakukan transformasi dengan metode minimum sebesar -20 dan maksimum sebesar 20 model seperti diatas dengan R-squared sebesar 0,195 dimana r-squared bernilai sama dengan data yang hanya diatasi missing valuenya.

1. **Mengatasi Missing Value, Outlier, dan Transformasi**

OLS Regression Results

==============================================================================

Dep. Variable: y4 R-squared: 0.907

Model: OLS Adj. R-squared: 0.907

Method: Least Squares F-statistic: 1.338e+05

Date: Sat, 09 Mar 2019 Prob (F-statistic): 0.00

Time: 16:14:00 Log-Likelihood: 98145.

No. Observations: 40995 AIC: -1.963e+05

Df Residuals: 40991 BIC: -1.962e+05

Df Model: 3

Covariance Type: nonrobust

==============================================================================

coef std err t P>|t| [0.025 0.975]

------------------------------------------------------------------------------

Intercept 6.4297 0.011 568.345 0.000 6.408 6.452

x4[0] 1.2148 0.015 82.526 0.000 1.186 1.244

x4[1] -1.0097 0.017 -58.838 0.000 -1.043 -0.976

x4[2] -6.3737 0.011 -556.330 0.000 -6.396 -6.351

==============================================================================

Omnibus: 1682.675 Durbin-Watson: 0.199

Prob(Omnibus): 0.000 Jarque-Bera (JB): 1897.744

Skew: -0.526 Prob(JB): 0.00

Kurtosis: 3.072 Cond. No. 298.

==============================================================================

Dari output diatas jika dilakukan pembuangan outlier, transformasi dengan di normalkan model seperti diatas dengan R-squared sebesar 0,907

KESIMPULAN

Hasil dari model terbaik di berikan oleh data dengan diatasi missing value, dibuang outlier, dan ditransformasikan. Karena pada data tersebut sudah dilakukan pre-processing data secara lengkap sehingga diperoleh nilai R-squared terbesar sebesar 0,907.