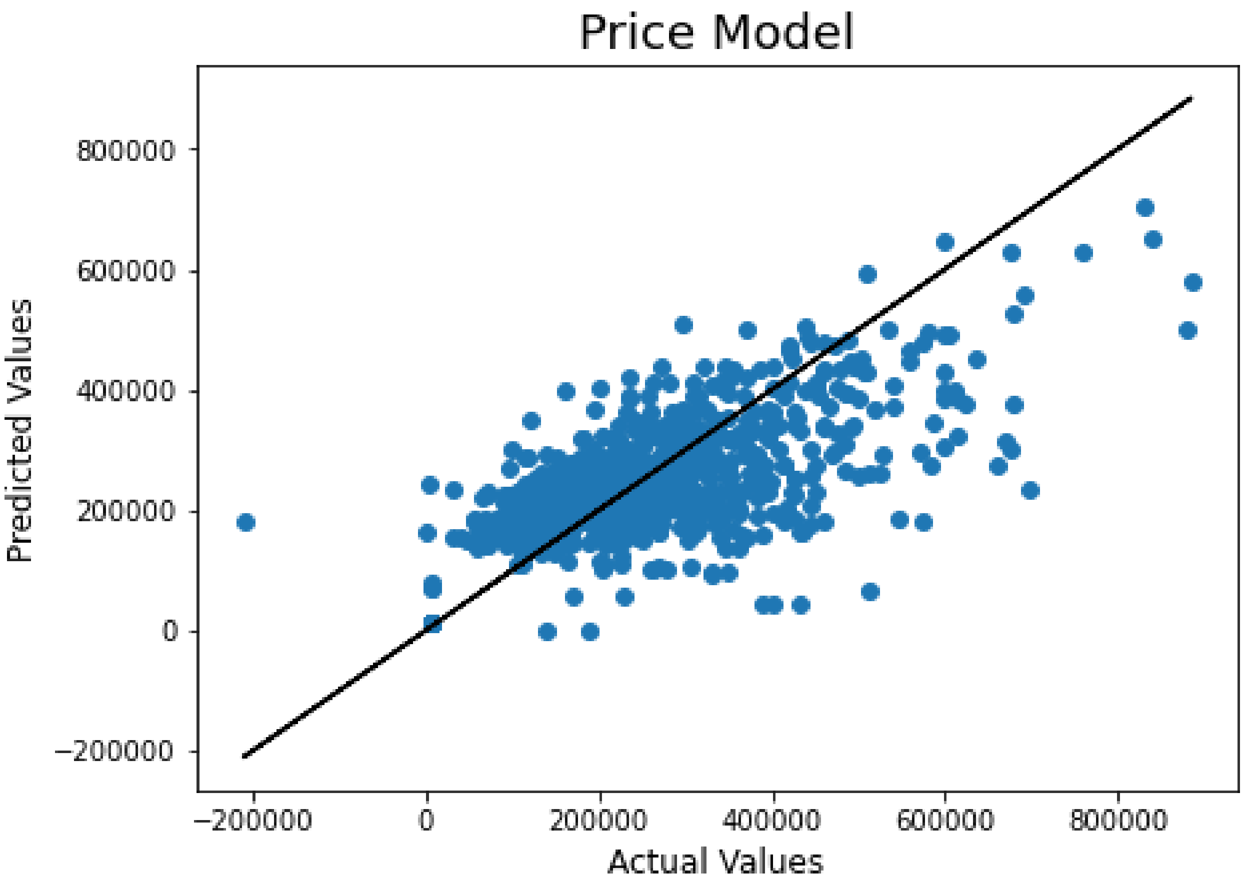
Sacramento Pricing Model

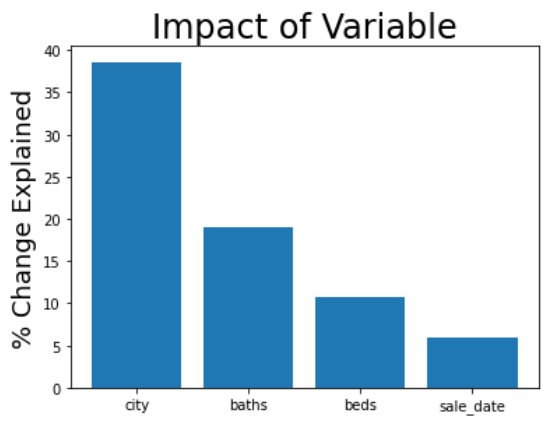


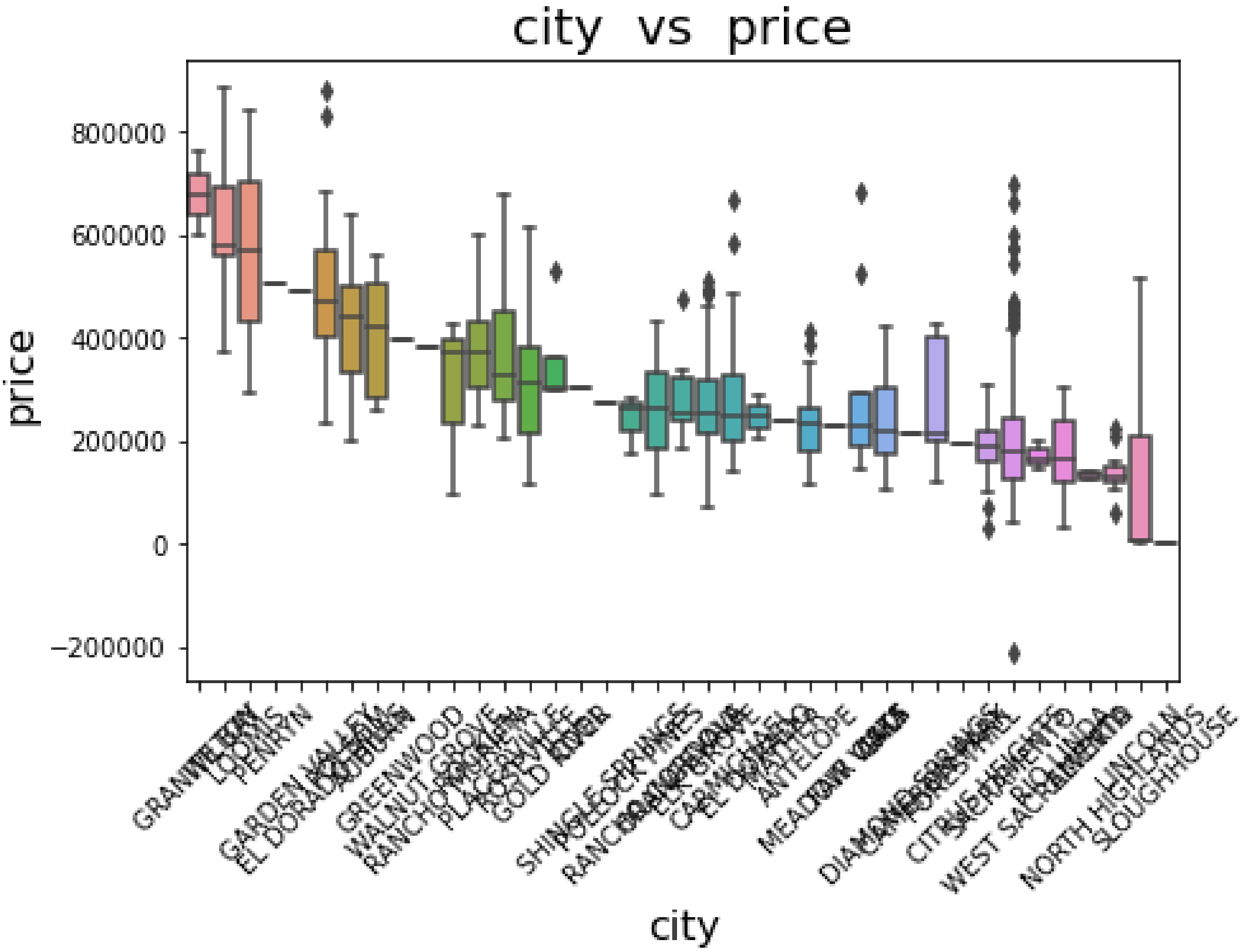
	feature	original_feature	engineering	type	r2_score
0	city_mode_imputed	city	mode_imputed	categorical	0.385468
1	baths_mean_imputed^2	baths	mean_imputed^2	numeric	0.189324
2	beds_mean_imputed^2	beds	mean_imputed^2	numeric	0.107575
3	sale_date_mode_imputed	sale_date	mode_imputed	categorical	0.059525

Main Model Summary

```
                            OLS Regression Results
=======================================================================================
Dep. Variable:                  price   R-squared (uncentered):                   0.867
Model:                            OLS   Adj. R-squared (uncentered):              0.866
Method:                 Least Squares   F-statistic:                              1593.
Date:                Fri, 14 Oct 2022   Prob (F-statistic):                        0.00
Time:                        00:52:28   Log-Likelihood:                         -12731.
No. Observations:                 985   AIC:                                  2.547e+04
Df Residuals:                     981   BIC:                                  2.549e+04
Df Model:                           4
Covariance Type:            nonrobust
==========================================================================================
                            coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------------------
city_mode_imputed         0.7063      0.033     21.109      0.000       0.641       0.772
baths_mean_imputed^2   1.254e+04   1456.901      8.608      0.000    9682.503    1.54e+04
beds_mean_imputed^2     793.8100    697.757      1.138      0.256    -575.458    2163.078
sale_date_mode_imputed    0.1265      0.038      3.358      0.001       0.053       0.200
==============================================================================
Omnibus:                      230.700   Durbin-Watson:                   0.916
Prob(Omnibus):                  0.000   Jarque-Bera (JB):              547.043
Skew:                           1.251   Prob(JB):                    1.63e-119
Kurtosis:                       5.659   Cond. No.                     1.53e+05
==============================================================================

Notes:
[1] R² is computed without centering (uncentered) since the model does not contain a constant.
[2] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[3] The condition number is large, 1.53e+05. This might indicate that there are
strong multicollinearity or other numerical problems.
```



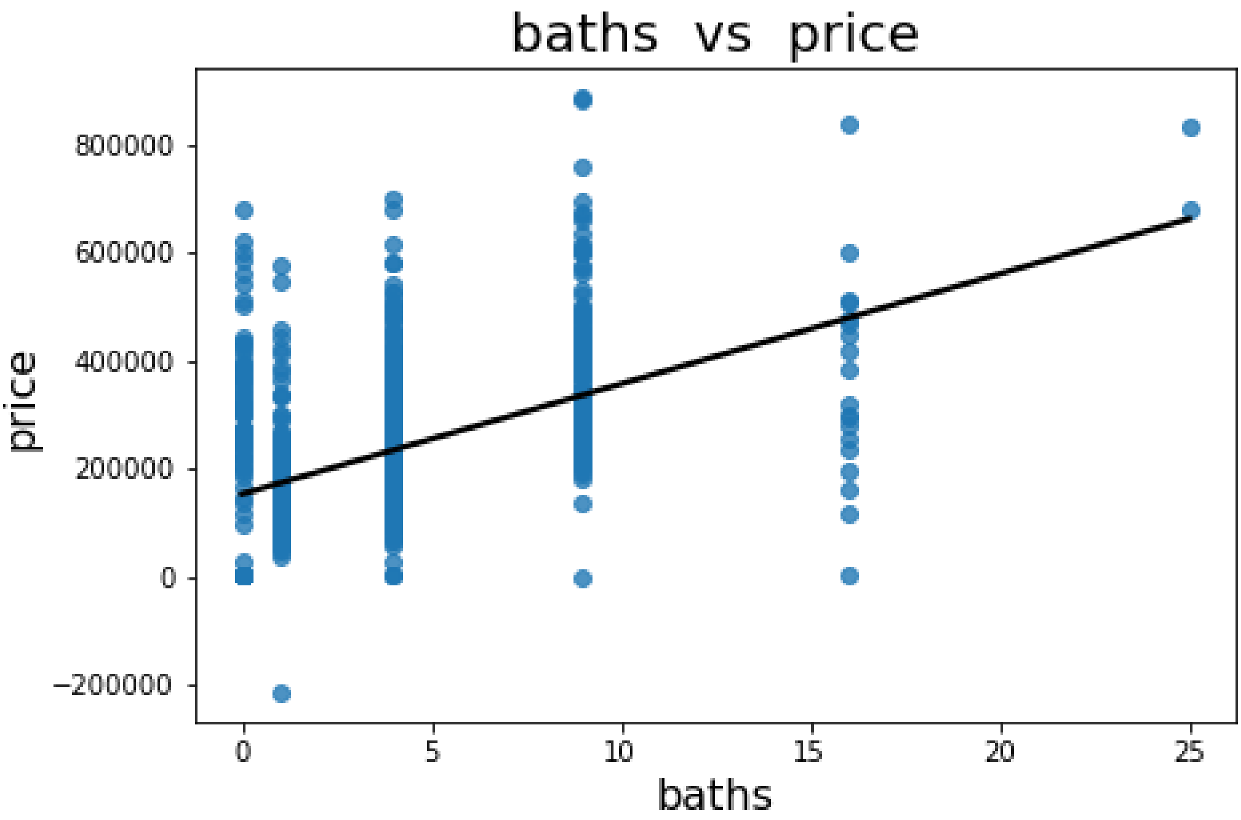


Variable Type: categorical

Engineering: mode_imputed

```
                                 OLS Regression Results
===================================================================================
Dep. Variable:      city_mode_imputed   R-squared (uncentered):
Model:                            OLS   Adj. R-squared (uncentered):
Method:                 Least Squares   F-statistic:
Date:                Fri, 14 Oct 2022   Prob (F-statistic):
Time:                        00:52:29   Log-Likelihood:                          -1
No. Observations:                 985   AIC:                                   2.53
Df Residuals:                     984   BIC:                                   2.54
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
price          0.7871      0.011     70.121      0.000       0.765       0.809
==============================================================================
Omnibus:                      179.068   Durbin-Watson:                   1.216
Prob(Omnibus):                  0.000   Jarque-Bera (JB):              457.693
Skew:                          -0.955   Prob(JB):                    4.10e-100
Kurtosis:                       5.740   Cond. No.                         1.00
==============================================================================

Notes:
[1] R² is computed without centering (uncentered) since the model does not contair
[2] Standard Errors assume that the covariance matrix of the errors is correctly s
```

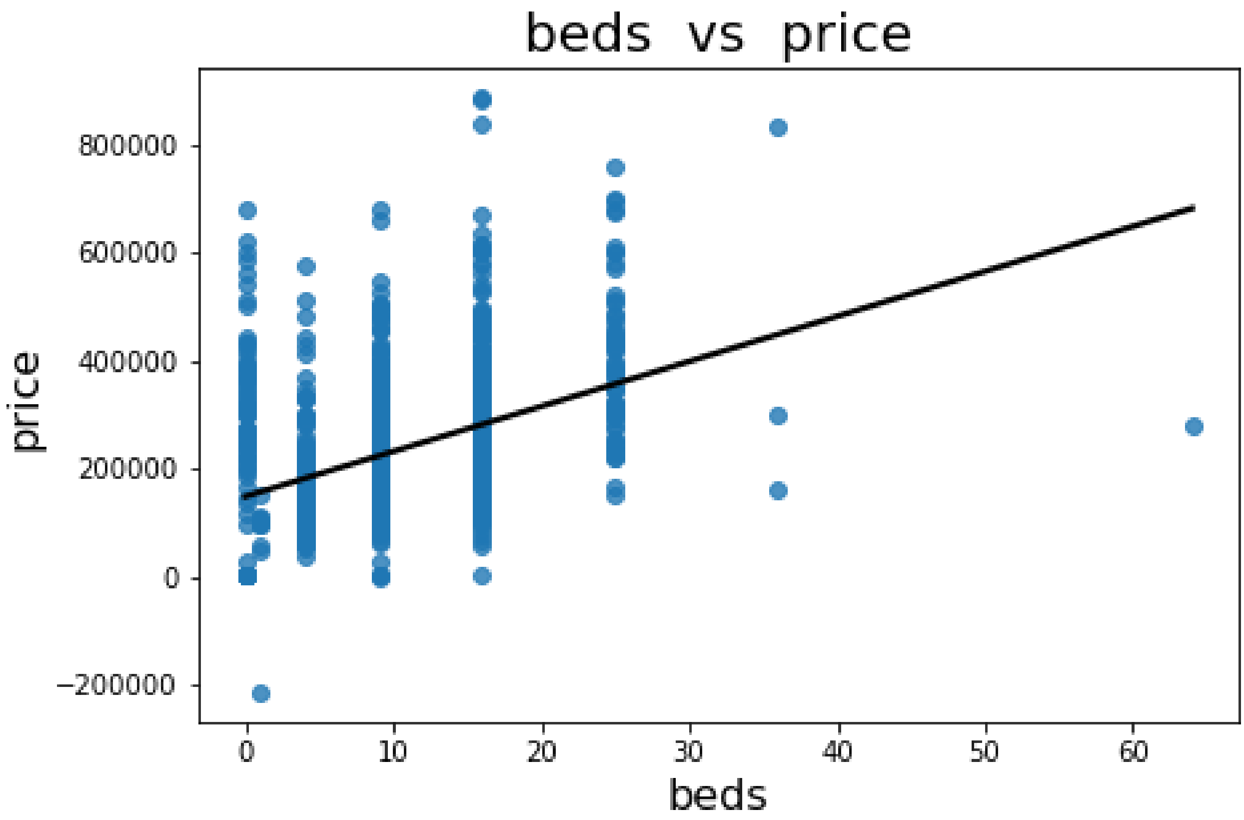


Variable Type: numeric

Engineering: mean_imputed^2

```
                                 OLS Regression Results
=====================================================================================
Dep. Variable:     baths_mean_imputed^2   R-squared (uncentered):
Model:                              OLS   Adj. R-squared (uncentered):
Method:                   Least Squares   F-statistic:
Date:                  Fri, 14 Oct 2022   Prob (F-statistic):                     3.
Time:                          00:52:30   Log-Likelihood:
No. Observations:                   985   AIC:
Df Residuals:                       984   BIC:
Df Model:                             1
Covariance Type:              nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
price       1.537e-05   3.44e-07     44.701      0.000    1.47e-05     1.6e-05
==============================================================================
Omnibus:                      168.890   Durbin-Watson:                   1.837
Prob(Omnibus):                  0.000   Jarque-Bera (JB):              884.689
Skew:                           0.676   Prob(JB):                    7.80e-193
Kurtosis:                       7.441   Cond. No.                         1.00
==============================================================================

Notes:
[1] R² is computed without centering (uncentered) since the model does not contair
[2] Standard Errors assume that the covariance matrix of the errors is correctly s
```

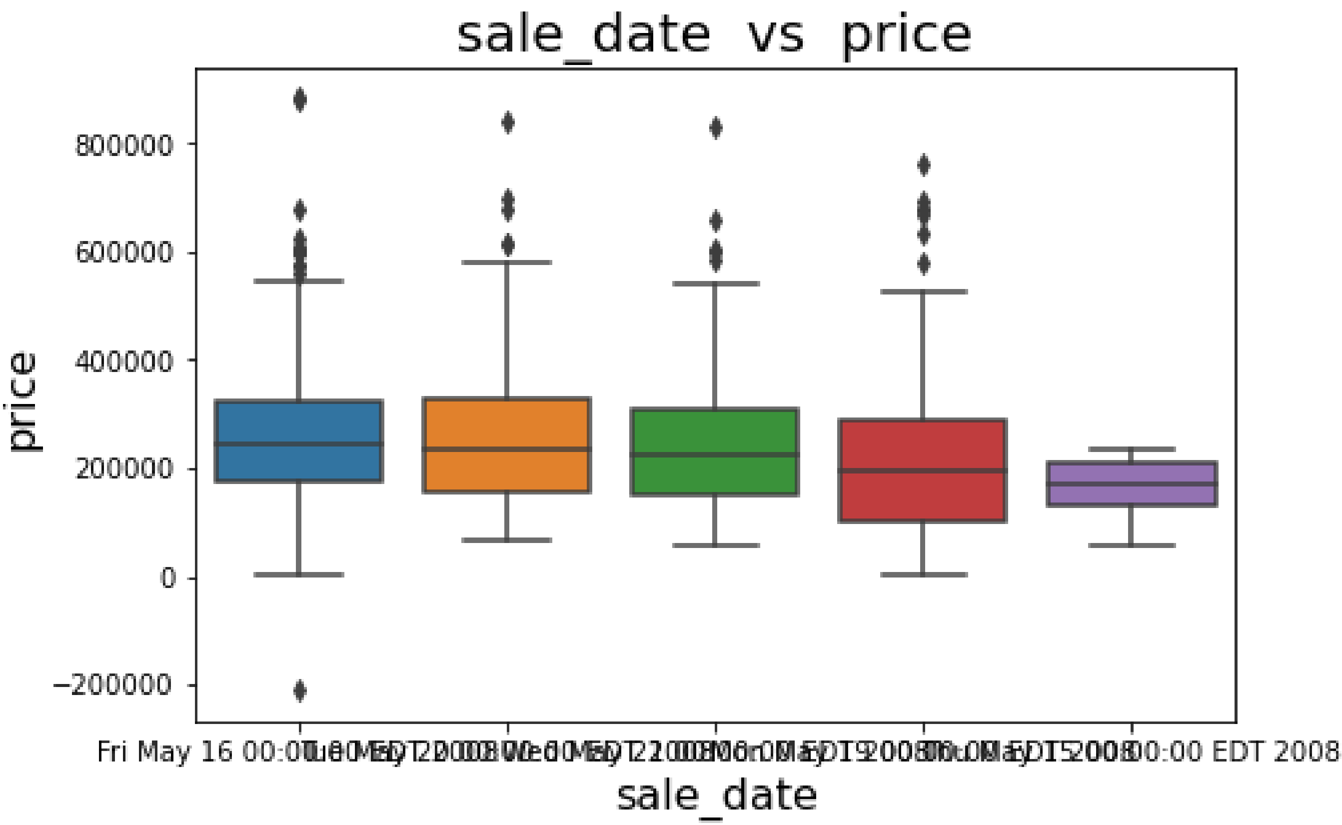


Variable Type: numeric

Engineering: mean_imputed^2

```
                                 OLS Regression Results
===================================================================================
Dep. Variable:     beds_mean_imputed^2   R-squared (uncentered):
Model:                             OLS   Adj. R-squared (uncentered):
Method:                  Least Squares   F-statistic:
Date:                 Fri, 14 Oct 2022   Prob (F-statistic):                    2.6
Time:                         00:52:30   Log-Likelihood:                         -
No. Observations:                  985   AIC:
Df Residuals:                      984   BIC:
Df Model:                            1
Covariance Type:             nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
price       3.725e-05   7.96e-07     46.787      0.000    3.57e-05    3.88e-05
==============================================================================
Omnibus:                      116.082   Durbin-Watson:                   1.529
Prob(Omnibus):                  0.000   Jarque-Bera (JB):             1028.385
Skew:                           0.033   Prob(JB):                    4.89e-224
Kurtosis:                       8.005   Cond. No.                         1.00
==============================================================================

Notes:
[1] R² is computed without centering (uncentered) since the model does not contair
[2] Standard Errors assume that the covariance matrix of the errors is correctly s
```



Variable Type: categorical

Engineering: mode_imputed

```
                            OLS Regression Results
=================================================================================
Dep. Variable:      sale_date_mode_imputed   R-squared (uncentered):
Model:                                 OLS   Adj. R-squared (uncentered):
Method:                      Least Squares   F-statistic:
Date:                     Fri, 14 Oct 2022   Prob (F-statistic):
Time:                             00:52:31   Log-Likelihood:
No. Observations:                      985   AIC:
Df Residuals:                          984   BIC:
Df Model:                                1
Covariance Type:                 nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
price          0.6984      0.013     54.845      0.000       0.673       0.723
==============================================================================
Omnibus:                      180.284   Durbin-Watson:                   0.127
Prob(Omnibus):                  0.000   Jarque-Bera (JB):              371.647
Skew:                          -1.043   Prob(JB):                     1.99e-81
Kurtosis:                       5.168   Cond. No.                         1.00
==============================================================================

Notes:
[1] R² is computed without centering (uncentered) since the model does not contair
[2] Standard Errors assume that the covariance matrix of the errors is correctly s
```