maXbox6

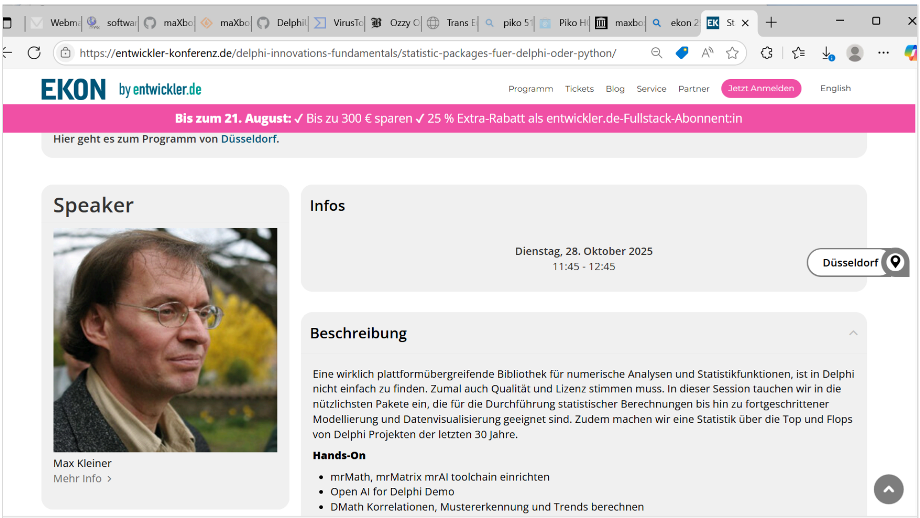
all about code blocks

About
EKON 30
maXbox
RockChain
TEE

recent posts

Statistics with Statsmodels
Statsmodels for Python and Delphi
Japan Religion
Hello Code
RockChain World!

about



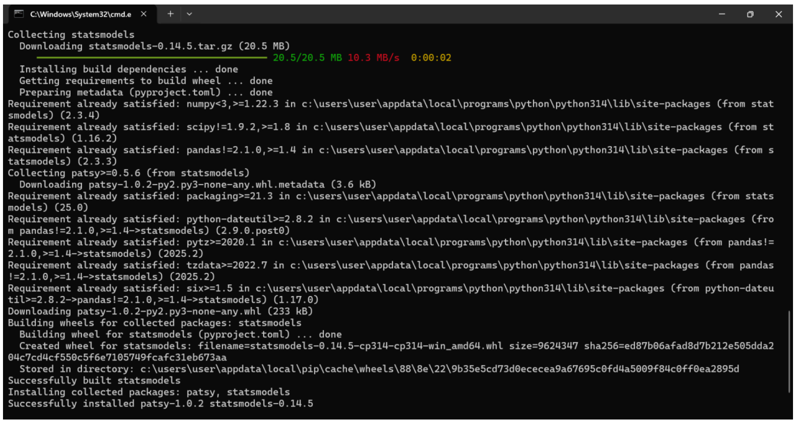
Statsmodels for Python and Delphi

November 21, 2025

Statsmodels is a Python library designed for statistical modeling, hypothesis testing, and data exploration with high reference. It provides a wide range of statistical models, including linear regression, time series analysis, and generalized linear models.

The library supports both formula-based modeling (similar to R) and direct use of NumPy arrays. First you have to install the library package:

py -3.14 -m pip install statsmodels



C:\maxbox\maxbox5\pascriptmaster2\pmaster2\restunits\Indy9\maxbox5\maxbox5>py -3.14 -m pip install statsmodels

The **OLS summary report is a detailed output that provides various metrics and statistics to help evaluate the model's performance and interpret its results**. Understanding each one can reveal valuable insights into your model's performance and accuracy. The summary table of the regression is given below for reference, providing detailed information on the model's performance, the significance of each variable, and other key statistics that help in interpreting the results. Here are the key components of the OLS summary:

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                Lottery   R-squared:                       0.414
Model:                            OLS   Adj. R-squared:                  0.392
Method:                 Least Squares   F-statistic:                     19.30
Date:                Thu, 26 Jun 2025   Prob (F-statistic):           1.47e-09
Time:                        17:35:22   Log-Likelihood:                -375.28
No. Observations:                  86   AIC:                             758.6
Df Residuals:                      82   BIC:                             768.4
Df Model:                           3
Covariance Type:            nonrobust
===============================================================================
```

OLS Results – Demo: 1417_statsmodels_64_delphi_python3.12.4debug130_EKON29.txt

Many statistical software options, like MATLAB, Minitab, Delphi, SPSS, and R, are available for regression analysis, this article focuses on using Python.

Ordinary Least Squares (OLS) is a method used to estimate the coefficients of a linear regression model, such as y = ax + b, where *a* is the slope and *b* is the intercept. The goal is to minimize the sum of squared residuals (differences between observed and predicted values).

Example: Calculating OLS Coefficients

```
import numpy as np

# Sample data
x = np.array([1, 2, 3, 4, 5]) # Independent variable
y = np.array([2.2, 2.8, 4.5, 3.7, 5.5]) # Dependent variable

# Adding a column of ones for the intercept term
X = np.vstack((np.ones(len(x)), x)).T

# OLS formula: β = (X'X)^(-1) X'y
beta = np.linalg.inv(X.T @ X) @ X.T @ y

print(f"Intercept (b): {beta[0]:.2f}, Slope (a): {beta[1]:.2f}")
```

Output:

```
Intercept (b): 1.31, Slope (a): 0.85
```

This means the regression line is approximately **y = 0.85x + 1.31**.

Example: Calculating OLS Coefficients

```
import numpy as np
# Sample data
x = np.array([1, 2, 3, 4, 5]) # Independent variable
y = np.array([2.2, 2.8, 4.5, 3.7, 5.5]) # Dependent variable

# Adding a column of ones for the intercept term
X = np.vstack((np.ones(len(x)), x)).T

# OLS formula: β = (X'X)^(-1) X'y

beta = np.linalg.inv(X.T @ X) @ X.T @ y
print(f"Intercept (b): {beta[0]:.2f}, Slope (a): {beta[1]:.2f
```

Output:

Intercept (b): 1.31, Slope (a): 0.85

This means the regression line is approximately y = 0.85x + 1.31.

Key Concepts

1. Intercept (b): Represents the value of *y* when *x = 0*.
2. Slope (a): Indicates the change in *y* for a one-unit increase in *x*.

The coefficients are derived using the formula: $\beta = (X^T X)^{-1} X^T y$, where:

- *X* is the design matrix (includes a column of ones for the intercept).
- *y* is the vector of observed values.

Important Considerations

- Assumptions: OLS assumes linearity, independence of errors, homoscedasticity (constant variance), and normally distributed errors.
- Multicollinearity: If predictors are highly correlated, it can affect coefficient estimates.
- Diagnostics: Use R-squared and residual analysis to evaluate model fit.

Advanced Example

Then you can build also graphic regression results with very small samples of code and the help of seaborn and matplotlib (you find comments in code):

Demo: 1417_statsmodels_64_delphi_python3.12.4debug130_EKON29.txt

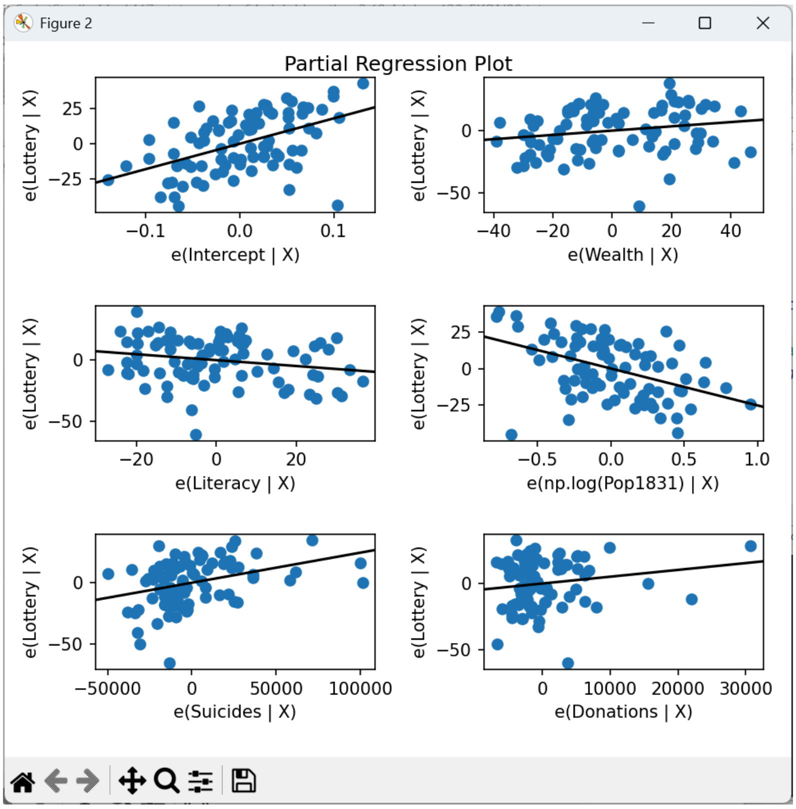
https://sourceforge.net/projects/maxbox5/files/EKON29/1417_statsmodels_64_delphi_python3.12.4debug130_EKON29.txt/download

Second multivariate with a higher R2 (Dependent Variable is Lottery):

```
//# Fit an OLS regression model
  Execstr('model = smf.ols("Lottery ~ Wealth+Literacy+np.log(
//# Display the summary of results
  execstr('print(model.summary())');
   {Notice that there is one missing observation in the Regio
             We eliminate it using a DataFrame method provide
  execstr('df = df.dropna()');
//execstr('fig, ax = plt.subplots(figsize=(12,8))');
  execstr('sm.graphics.plot_partregress("Lottery","Wealth",["

  execstr('sm.graphics.plot_partregress_grid(model)');
//https://github.com/statsmodels/statsmodels/blob/main/exampl
  ExecStr('plt.show()');
```

```
 Execstr('model=smf.ols("Lottery~ Wealth+Literacy+np.log(Pop1831
)+Suicides+Donations",data=df).fit()');
```



Partial Plots from Guerry Moral Statistic Dataset

R-squared measures the proportion of the variance in the dependent variable that is explained by the independent variables. It ranges from 0 to 1, where 1 indicates that the model explains all the variance.

OLS Regression Results

Dep. Variable: Lottery R-squared: 0.414
Model: OLS Adj. R-squared: 0.392
Method: Least Squares F-statistic: 19.30
Date: Fri, 21 Nov 2025 Prob (F-statistic): 1.47e-09
Time: 13:37:22 Log-Likelihood: -375.28
No. Observations: 86 AIC: 758.6
Df Residuals: 82 BIC: 768.4
Df Model: 3

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                Lottery   R-squared:                       0.492
Model:                            OLS   Adj. R-squared:                  0.461
Method:                 Least Squares   F-statistic:                     15.51
Date:                Thu, 02 Oct 2025   Prob (F-statistic):           1.20e-10
Time:                        17:54:49   Log-Likelihood:                -369.10
No. Observations:                  86   AIC:                             750.2
Df Residuals:                      80   BIC:                             764.9
Df Model:                           5
Covariance Type:            nonrobust
===============================================================================
                  coef    std err          t      P>|t|      [0.025      0.975]
-------------------------------------------------------------------------------
Intercept       180.2691     35.830      5.031     0.000     108.965     251.573
Wealth            0.1715      0.094      1.833     0.070      -0.015       0.358
Literacy         -0.2409      0.126     -1.911     0.060      -0.492       0.010
np.log(Pop1831) -25.1121      5.699     -4.407     0.000     -36.453     -13.771
Suicides          0.0002   7.37e-05      3.323     0.001    9.82e-05       0.000
Donations         0.0005      0.000      1.490     0.140      -0.000       0.001
==============================================================================
Omnibus:                        6.163   Durbin-Watson:                   1.665
Prob(Omnibus):                  0.046   Jarque-Bera (JB):                5.537
Skew:                          -0.594   Prob(JB):                       0.0628
Kurtosis:                       3.363   Cond. No.                     8.85e+05
==============================================================================
```

This is R2 of 0.492 R-Squared

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 1.13e+03. This might indicate that there are strong multicollinearity or other numerical problems.
dept Region Department Crime_pers Crime_prop Literacy Donations Infants Suicides MainCity Wealth Commerce Clergy Crime_parents Infanticide Donation_clergy Lottery Desertion Instruction Prostitutes Distance Area Pop1831

The F-statistic is used to test the overall significance of the model. The null hypothesis is that all coefficients (except the intercept) are zero, meaning the model does not explain any variance in the dependent variable. The p-value associated with the F-statistic indicates the probability of observing the F-statistic (or a more extreme value) if the null hypothesis were true. A small p-value (typically less than 0.05) indicates that the model is statistically significant, meaning at least one of the independent variables has a significant effect on the dependent variable Lottery.

Statsmodels Code

- We download the Guerry dataset, a collection of historical data used in support of Andre-Michel Guerry's 1833 Essay on the Moral Statistics of France. The data set is hosted
- online in comma-separated values format (CSV) by the Rdatasets repository. We could download the file locally and then load it using read_csv, but pandas takes care of all of this automatically for us:

Execstr('model = smf.ols("Lottery ~ Wealth + Literacy + np.log(Pop1831)", data=df).fit()');

```
Execstring('import statsmodels.api as sm; import numpy as np');

Execstr('model = smf.ols("Lottery ~ Wealth + Literacy + np.log(Pop1831)",
execstr('print(model.summary())');
execstr('sm.graphics.plot_partregress_grid(model)');
```

[Download 1417_statsmodels_64_delphi_python3.12.4debug130_EKON29.txt (maXbox5)]

[Interpreting the results of Linear Regression using OLS Summary – GeeksforGeeks]

[EKON 29 Statistic Packages Slides | maXbox]

Posted in Code, Data Science, maXbox, Python

One response to "Statsmodels for Python and Delphi"

Statistics with Statsmodels – maXbox6
November 21, 2025 Edit

[...] Statsmodels for Python and Delphi [...]

★ Like

Reply

Leave a comment

← Japan Religion

Statistics with Statsmodels →

Blog at WordPress.com.

maXbox4exe Digitally signed by maXbox4exe
Date: 2025.11.21 16:33:01 +01'00'