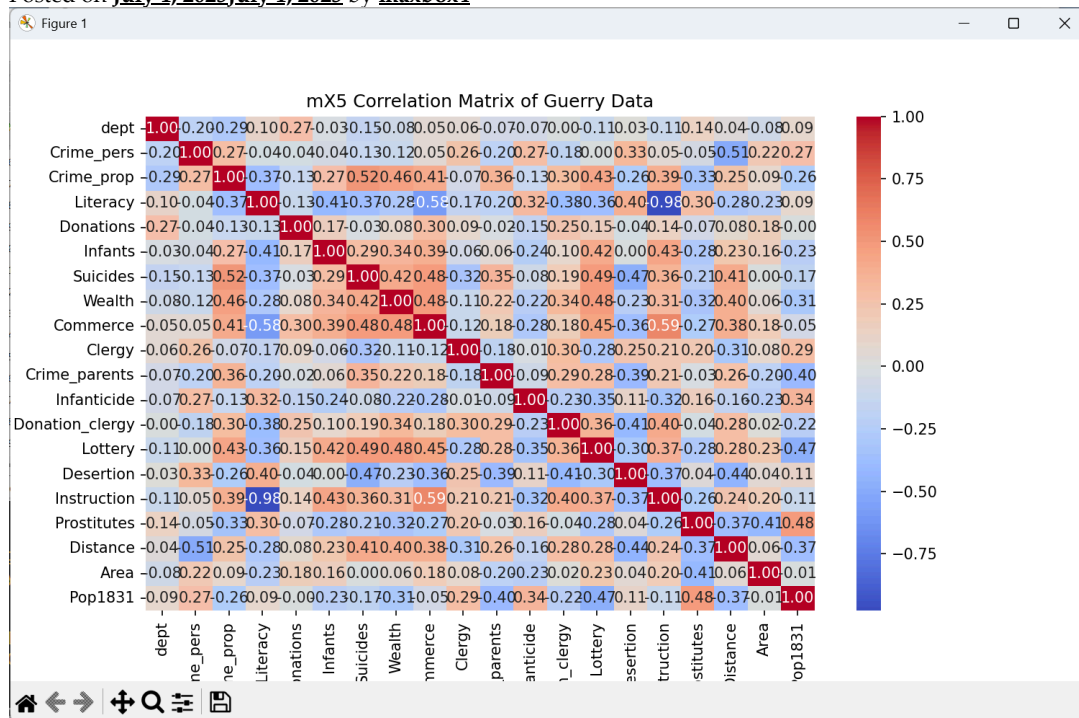


# maXbox

## EKON 29

Posted on [July 1, 2025](#) [July 1, 2025](#) by [maxbox4](#)



## Statistic Packages for Delphi and Python

A truly cross-platform library for numerical analysis and statistical functions isn't easy to find in Delphi. Especially since quality and licensing are also important. In this session, we'll delve into the most useful packages and libraries suitable for performing statistical calculations, all the way up to advanced modeling and data visualization. We'll also compile a statistics on the top and flops of Delphi projects over the past 30 years.

<https://entwickler-konferenz.de/delphi-innovations-fundamentals/statistic-packages-fuer-delphi-oder-python> (<https://entwickler-konferenz.de/delphi-innovations-fundamentals/statistic-packages-fuer-delphi-oder-python>).

### Hands-On

- Setting up the mrMath and mrMatrix mrAI toolchain
- Open AI for Delphi Demo
- Calculating DMath correlations, pattern recognition, and trends
- Introducing the 7 well-known statistical methods
- Demonstrating the 5 most important chart types (Bar Chart, Scatter Plot, Histogram, Box Plot and Correlationmatrix)
- Setting up an energy storage time series AGSI project statistics
- Descriptive statistics with the reference dataset for morale statistics (Guerry, "HistData")
- Data Science Tutorial AGSI or Guerry

### Contents

- Overview of the statistical packages with Delphi and Python (P4D).
- Configuration and spec features of DMath, SKLearn, Statsmodels, & DataLab library
- Troubleshooting: Typical miscalculations and their solutions with cleaned data
- First steps in implementing regression, cluster analysis, and correlation matrix

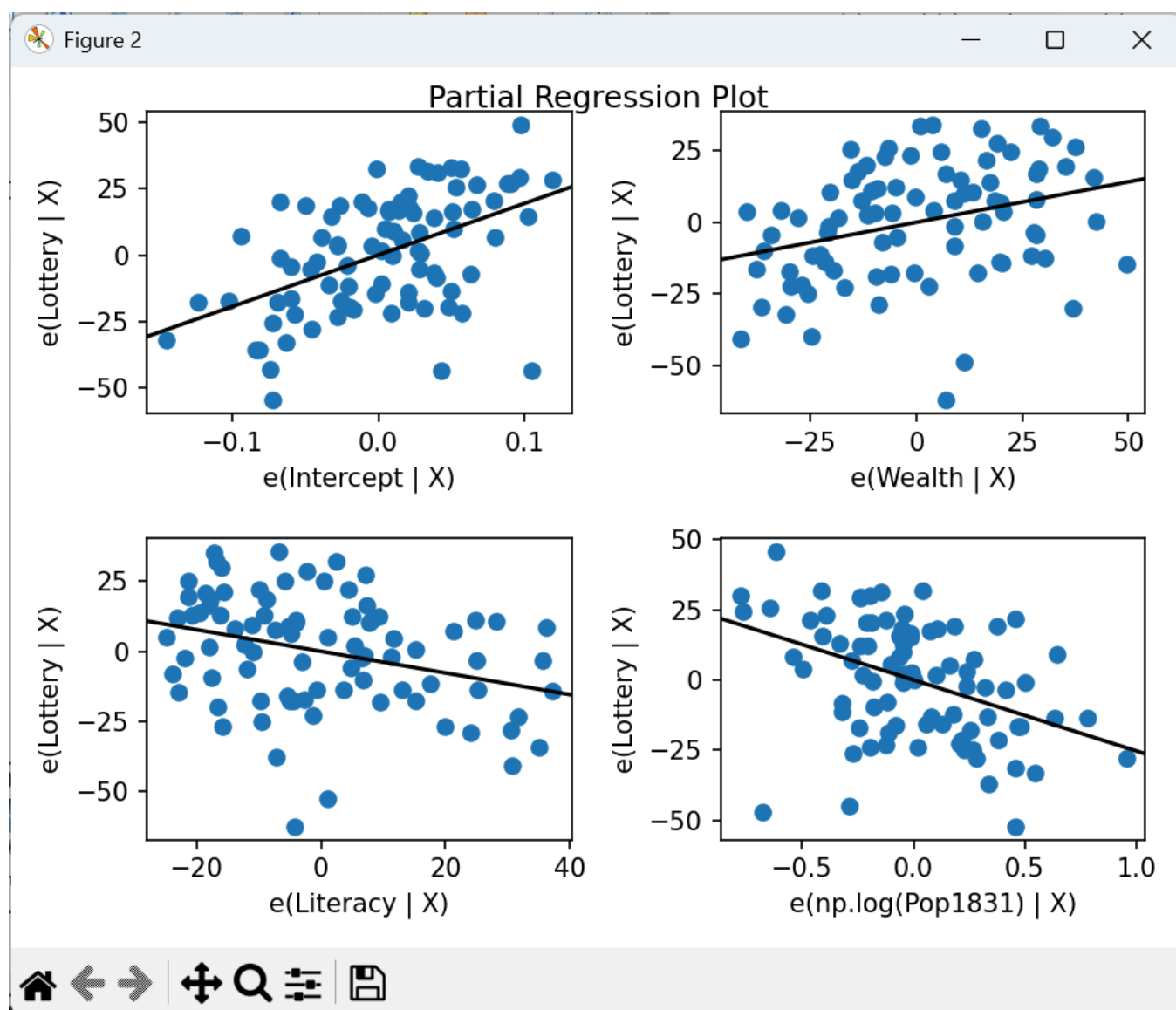
We download the Guerry dataset, a collection of historical data used in support of Andre-Michel Guerry's 1833 Essay on the Moral Statistics of France. The data set is hosted online in comma-separated values format (CSV) by the Rdatasets repository. We could download the file locally and then load it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

```

=====
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
=====
                    coef    std err          t      P>|t|    [0.025    0.975]
-----
Intercept          194.3951    37.768      5.147      0.000    119.263    269.527
Wealth              0.2820     0.093      3.024      0.003      0.097      0.468
Literacy           -0.3840     0.127     -3.033      0.003     -0.636     -0.132
np.log(Pop1831)   -25.2363     6.047     -4.174      0.000    -37.265    -13.207
=====
Omnibus:             7.602    Durbin-Watson:          1.890
Prob(Omnibus):       0.022    Jarque-Bera (JB):        7.051
Skew:               -0.651    Prob(JB):                0.0294
Kurtosis:            3.524    Cond. No.                1.13e+03
=====

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.
mX5 executed: 26/06/2025 17:35:23 Runtime: 0:0:47.680 Memload: 58% use

```



## Partial Regression Plot Grid

```

1 Execstring('import statsmodels.api as sm; import numpy as np');
2
3 Execstr('model = smf.ols("Lottery ~ Wealth + Literacy + np.log(Pop1831)", data=df).fit()');
4 execstr('print(model.summary())');
5 execstr('sm.graphics.plot_partregress_grid(model)');

```

**Statsmodels** is a Python **library** designed **for** statistical modeling, hypothesis testing, **and** data exploration. 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.

The **data science tutorial** explains the so called AGSI data storage **and** his visualization **of** the timeline. AGSI is the Aggregated Gas Storage Inventory **and** offers you the possibility **to** be kept up **to** date whenever a new service announcement **or** update from one **of** our data providers **is** posted **on** the website.

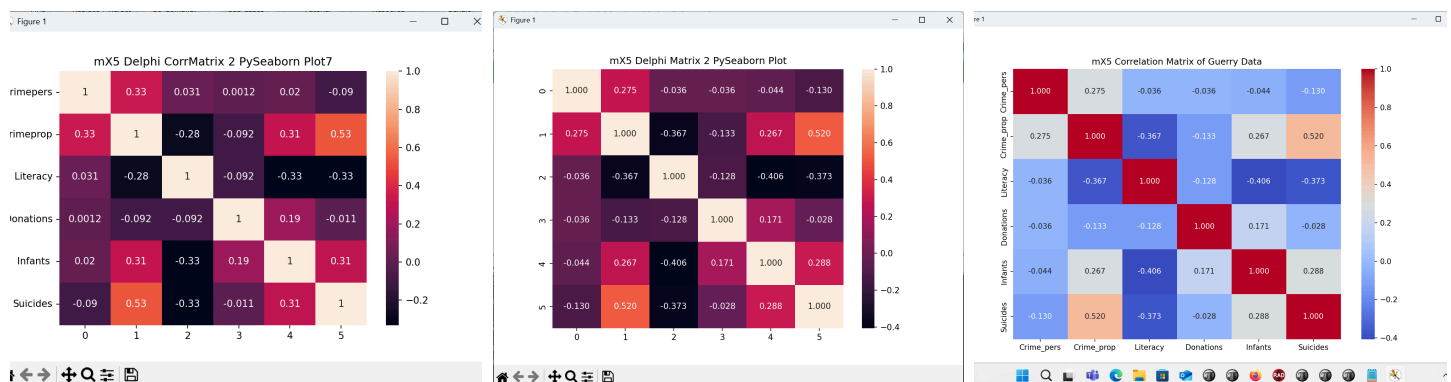
Data representation **of** gas in storage **as** a timeline AGSI dataset.

<https://blogs.embarcadero.com/why-a-data-scientist-chooses-delphi-for-powerful-real-world-visualizations/>  
[\(https://blogs.embarcadero.com/why-a-data-scientist-chooses-delphi-for-powerful-real-world-visualizations/\)](https://blogs.embarcadero.com/why-a-data-scientist-chooses-delphi-for-powerful-real-world-visualizations/)

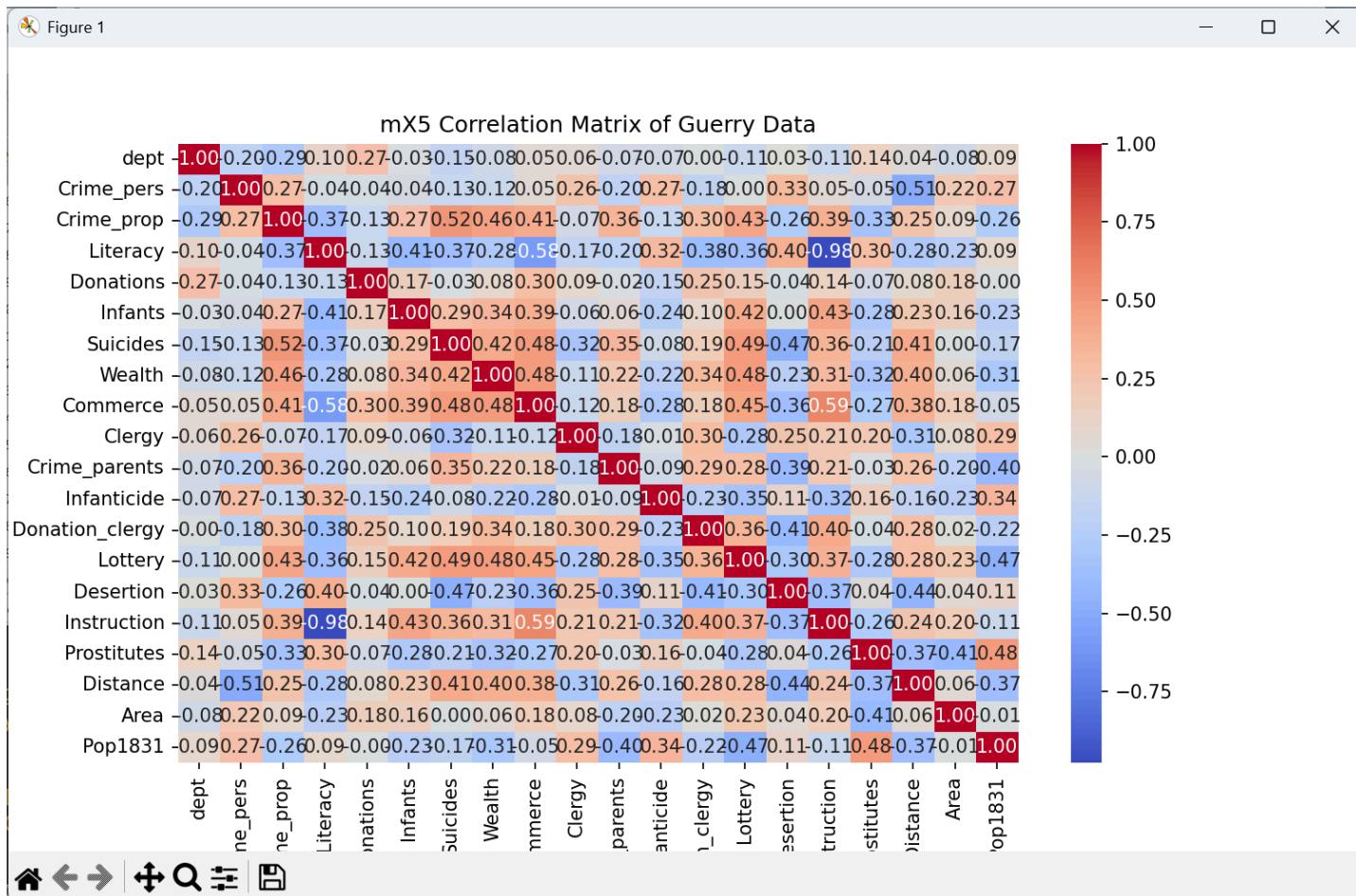
Scikitlearns model.score(X,y) calculation works **on** coefficient **of** determination i.e  $R^2$  **is** a simple **function** that takes model.score= (X\_test,y\_test). It doesn't require y\_predicted value **to** be supplied externally **to** calculate the score **for** you, rather it calculates y\_predicted internally **and** uses it **in** the calculations.

- Mean **Function**: Calculates the average **of** an **array**.
- Covariance **Function**: Computes the covariance between two arrays.
- Correlation: **Uses** covariance **and** standard deviations for the correlation coefficient.
- ComputeCorrelationMatrix: Iterates through all variable pairs **to** compute corr-matrix.
- PrintMatrix: Outputs the matrix **to** the console.

Delphi does **not** have a built-in correlation matrix **function** with a heatmap too, but you can implement one using standard math operations.



Seaborn Heatmap



Guerry Dataset – from a csv file to a dataframe

The **mrMath**, mrStats, mrMatrix, mrImgUtils package includes:

- Standard Fisher LDA classifier
- Robust (and Fast Robust) version of this classifier
- Incremental (and Robust) Fisher LDA classifier learning.
- Support Vector Machines (least squares and lagrangian learning)
- Naive Bayes
- Simple Decision stumps
- Radial basis function
- C4.5 Decision trees.
- K-means
- Ensemble classifiers: AdaBoost, Gentle Boost, Bagging
- Simple feed forward Neural Nets

The screenshot shows the maXbox5 64-bit ScriptStudio window. The top menu bar includes File, Program, Options, View, Debug, Output, and Help. Below the menu is a toolbar with icons for Load, Find, Replace / Refact, Go Compile!, and Use Cases. The main editor displays a Python script with the following code:

```

232     println(eng.evalStr('pendulum.datetime(2022, 1, 5)'));
233     ## Load sample dataset
234     eng.Execstring('data = sm.datasets.get_rdataset("Guerry", "HistData").data');
235     ##println(eng.evalStr('data.describe()'));
236     println(eng.evalStr('data.info()'));
237     ## Fit an OLS regression model
238     eng.Execstring('model = smf.ols("Lottery ~ Literacy + np.log(Pop1831)", data=data).fit()');
239     ## Display the summary of results
240     eng.execstr('print(model.summary())');
241     ##println(eng.evalStr('(output.getvalue().strip())'));
242     println(StringReplace(eng.evalStr('(output.getvalue().strip())'),
243                        LF, CRLF, [rfReplaceAll]));
244
245 except

```

The console output shows the results of the script execution, including the OLS regression summary:

```

=====
Dep. Variable:          Lottery    R-squared:                0.348
Model:                  OLS        Adj. R-squared:           0.333
Method:                 Least Squares    F-statistic:             22.20
Date:                  Thu, 26 Jun 2025    Prob (F-statistic):      1.90e-08
Time:                  16:12:59          Log-Likelihood:          -379.82
No. Observations:      86              AIC:                    765.6
Df Residuals:          83              BIC:                    773.0
Df Model:              2
Covariance Type:       nonrobust
=====

```

	coef	std err	t	P> t	[0.025	0.975]
Intercept	246.4341	35.233	6.995	0.000	176.358	316.510
Literacy	-0.4889	0.128	-3.832	0.000	-0.743	-0.235
np.log(Pop1831)	-31.3114	5.977	-5.239	0.000	-43.199	-19.424

The status bar at the bottom indicates the file path: maXbox5 C:\maxbox\maxbox5\examples\1190\_Continued\_fraction64\_python3.12.4debug30.txt, the current time: Ct:26/06/2025 16:12:58, memory usage: Mem:67%, runtime: Rtime:0:0:2.991, and thread count: Thr:32.

maX5.2 with Statsmodels Console Output

Learn how to **install statsmodels**, a Python package for statistical modeling, using Anaconda, PyPI, source or development version.

The screenshot shows a Windows Command Prompt window with the following output:

```

C:\Windows\System32\cmd.e  x  +  v

Collecting statsmodels
  Downloading statsmodels-0.14.4-cp313-cp313-win_amd64.whl.metadata (9.5 kB)
Requirement already satisfied: numpy<3,>=1.22.3 in c:\users\user\appdata\local\programs\python\python313\lib\site-packages (from statsmodels) (2.2.5)
Requirement already satisfied: scipy!=1.9.2,>=1.8 in c:\users\user\appdata\local\programs\python\python313\lib\site-packages (from statsmodels) (1.15.3)
Requirement already satisfied: pandas!=2.1.0,>=1.4 in c:\users\user\appdata\local\programs\python\python313\lib\site-packages (from statsmodels) (2.2.3)
Collecting patsy>=0.5.6 (from statsmodels)
  Using cached patsy-1.0.1-py2.py3-none-any.whl.metadata (3.3 kB)
Requirement already satisfied: packaging>=21.3 in c:\users\user\appdata\local\programs\python\python313\lib\site-packages (from statsmodels) (25.0)
Requirement already satisfied: python-dateutil>=2.8.2 in c:\users\user\appdata\local\programs\python\python313\lib\site-packages (from pandas!=2.1.0,>=1.4->statsmodels) (2.9.0.post0)
Requirement already satisfied: pytz>=2020.1 in c:\users\user\appdata\local\programs\python\python313\lib\site-packages (from pandas!=2.1.0,>=1.4->statsmodels) (2025.2)
Requirement already satisfied: tzdata>=2022.7 in c:\users\user\appdata\local\programs\python\python313\lib\site-packages (from pandas!=2.1.0,>=1.4->statsmodels) (2025.2)
Requirement already satisfied: six>=1.5 in c:\users\user\appdata\local\programs\python\python313\lib\site-packages (from python-dateutil>=2.8.2->pandas!=2.1.0,>=1.4->statsmodels) (1.17.0)
Downloading statsmodels-0.14.4-cp313-cp313-win_amd64.whl (9.8 MB)
9.8/9.8 MB 8.7 MB/s eta 0:00:00
Using cached patsy-1.0.1-py2.py3-none-any.whl (232 kB)
Installing collected packages: patsy, statsmodels
Successfully installed patsy-1.0.1 statsmodels-0.14.4

[notice] A new release of pip is available: 25.0.1 -> 25.1.1
[notice] To update, run: python.exe -m pip install --upgrade pip

C:\maxbox\maxbox5\pascrptmaster2\pmaster2\restunits\Indy9\maxbox5\maxbox5>exit

```

Statsmodels Py 3.13.4

You can compute a correlation matrix **in** Delphi by iterating over all pairs of variables, extracting columns, **and** applying the Pearson correlation formula. **For** more advanced matrix operations **or** large datasets, consider using a Delphi matrix **library**.

```

1  procedure ComputeCorrelationMatrix(const Data: DMatrix; var CorrMatrix: DMatrix);
2  var
3      i, j, k, nVars, nObs: Integer;
4      colI, colJ: array of Double;
5  begin
6      nObs:= Length(Data);
7      nVars:= Length(Data[0]);
8      //SetLength(CorrMatrix, nVars, nVars);
9      SetMatrixLength(corrMatrix, nvars, nvars);
10     for i:= 0 to nVars-1 do begin
11         SetLength(colI, nObs);
12         for j:= 0 to nObs - 1 do
13             colI[j]:= Data[j][i];
14         for j:= i to nVars-1 do begin
15             SetLength(colJ, nObs);
16             for k:= 0 to nObs - 1 do
17                 colJ[k]:= Data[k][j];
18             CorrMatrix[i][j]:= PearsonCorrelation(colI, colJ);
19             CorrMatrix[j][i]:= CorrMatrix[i][j]; // Matrix is symmetric
20         end;
21     end;
22 end;

```

To transform the CSV data from file to matrix and dataframe you need 4 steps:

```

1  S:= TStringList.Create;
2  try
3      //S.StrictDelimiter := True;
4      S.LineBreak := #10;
5      //S1.Delimiter := ',';
6      s.loadfromfile(exepath+'\examples\1417_export_dataframe.csv');
7      writ('size: '+itoa(s.count));
8      SetMatrixLength(mData, 86, 6);
9      TStringListToMatrix(s, mData);

```

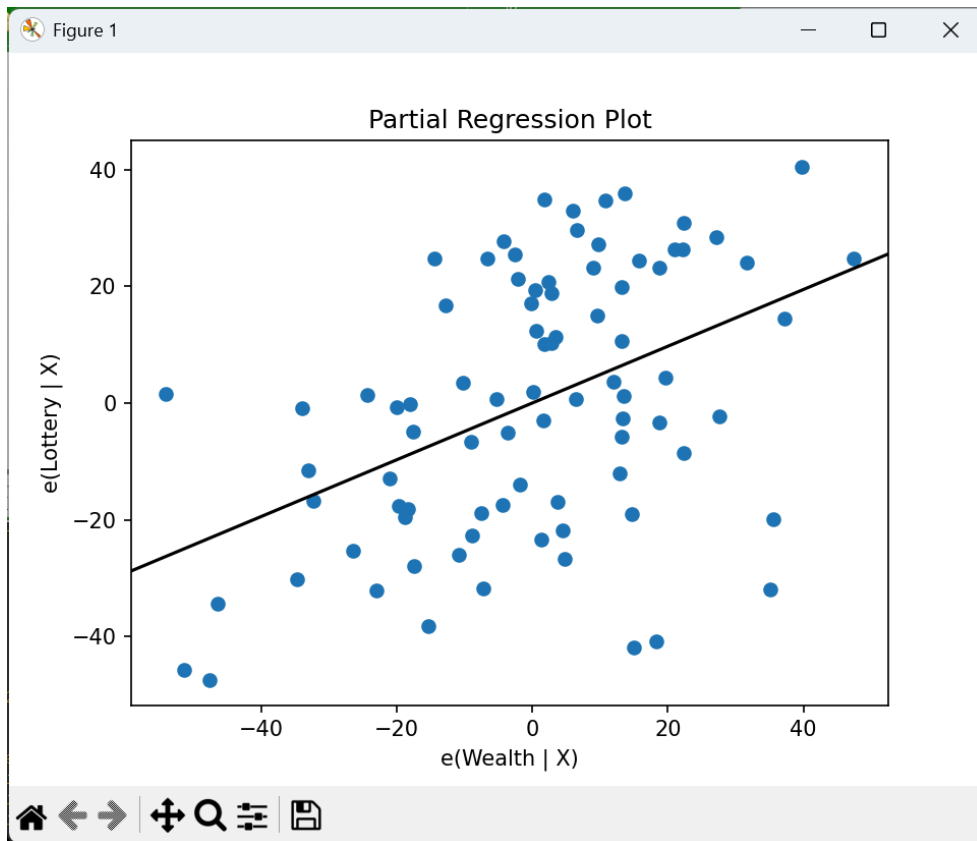
<b>Input:</b>	<b>Parsing:</b>	<b>Matrix Population:</b>	<b>Output:</b>
The StringList contains rows of data as strings, separated commas.	Each string is split into columns using CommaText of a temporary TStringList.	The parsed values are stored in a 2D array ( Matrix ).	The matrix is printed to verify the conversion.

```

1  procedure TStringListToMatrix(strList: TStringList; var matrix: DMatrix);
2  var i, j: Integer;
3      RowData: TStringList;
4  begin
5      if strList.Count = 0 then Exit;
6      // Create a temporary TStringList to parse each row
7      RowData:= TStringList.Create;
8      try
9          RowData.Delimiter:= ','; // Assuming comma-separated values
10         RowData.StrictDelimiter:= True;
11         //RowData.commatext
12         // Resize matrix to match the TStringList dimensions
13         SetLength(matrix, strList.Count);
14         for i := 1 to strList.Count - 1 do begin
15             RowData.DelimitedText:= strList[i];
16             //writ('debug '+itoa(rowdata.count));
17             SetLength(Matrix[i], RowData.Count);
18             for j:= 4 to RowData.Count - 15 do begin //slice 4-9
19                 Matrix[i][j-4]:= strtofloat(RowData[j]);
20                 //writ('debug '+flots(matrix[i][j]));
21             end;
22         end;
23     finally
24         RowData.Free;
25     end;
26 end;

```





Partial Regression Plot with Statsmodels

debug: 208-RuntimeError: CPU dispatcher tracer already initlized 865 err:20

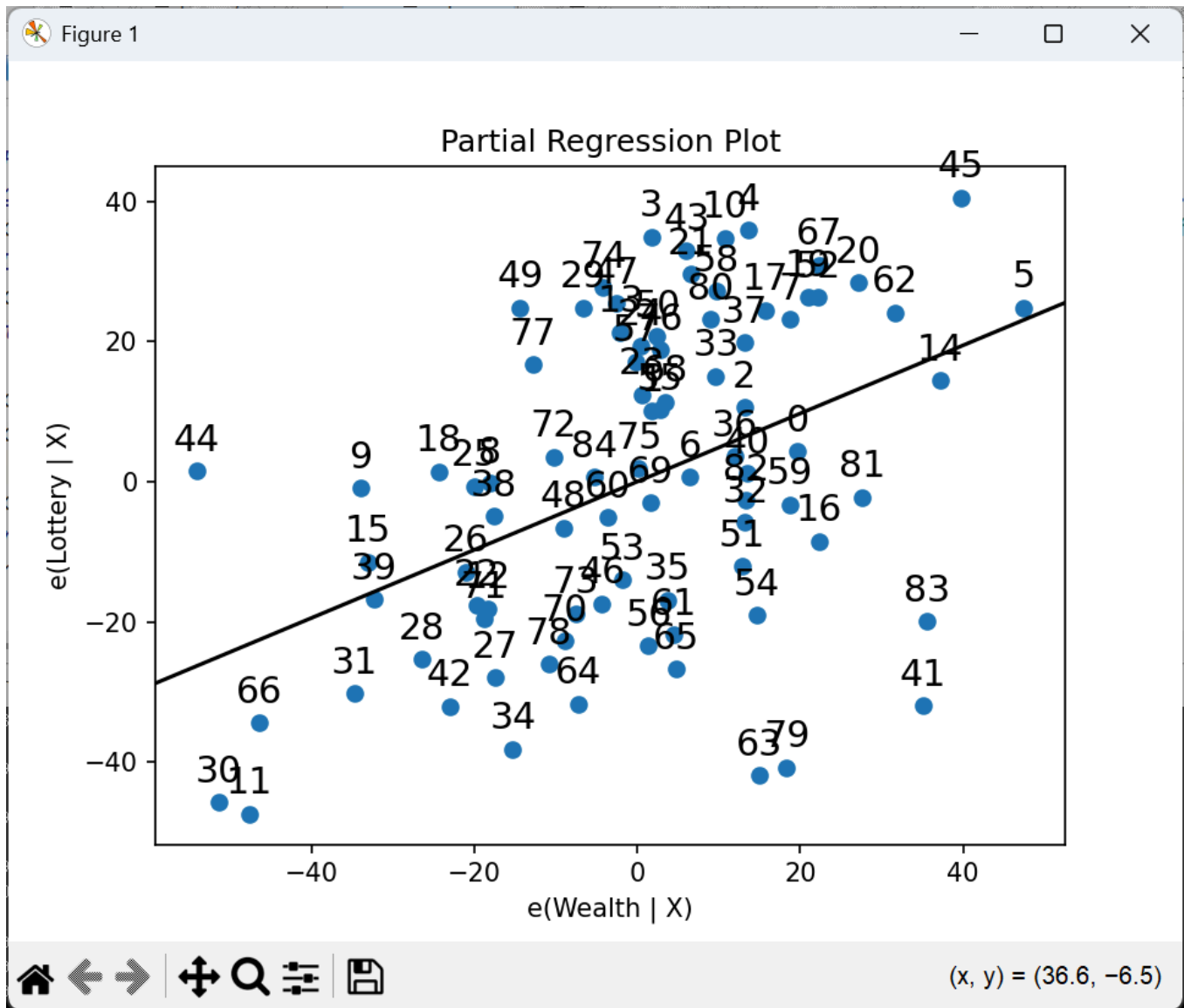
debug: 209-RuntimeError: CPU dispatcher tracer already initlized 865 err:20

Exception: RuntimeError: CPU dispatcher tracer already initlized at 865.3134

```

1  1  // Fit an OLS regression model
2  2  //eng.Execstring('model = smf.ols("Lottery ~ Literacy + np.log(Pop1831)", data=data).fit()');
3  3  Execstr('model = smf.ols("Lottery ~ Wealth + Literacy + np.log(Pop1831)", data=df).fit()');
4  4  // Display the summary of results
5  5  execstr('print(model.summary())');
6  6  {Notice that there is one missing observation in the Region column.
7  7  We eliminate it using a DataFrame method provided by pandas:}
8  8  execstr('df = df.dropna()');
9  9  execstr('sm.graphics.plot_partregress("Lottery", "Wealth", ["Region", "Distance"], '+
10 10                                     'data=df, obs_labels=False)');
11 11  ExecStr('plt.show()');

```



sample or observation tracer obs\_labels=True

Scripts at:

[https://sourceforge.net/projects/maxbox5/files/EKON29/1385\\_DCorrelation3SeabornPyCompare2\\_uc.txt/download](https://sourceforge.net/projects/maxbox5/files/EKON29/1385_DCorrelation3SeabornPyCompare2_uc.txt/download)  
[https://sourceforge.net/projects/maxbox5/files/EKON29/1385\\_DCorrelation3SeabornPyCompare2\\_uc.txt/download](https://sourceforge.net/projects/maxbox5/files/EKON29/1385_DCorrelation3SeabornPyCompare2_uc.txt/download)

[https://sourceforge.net/projects/maxbox5/files/examples/1417\\_statsmodels\\_64\\_delphi\\_python3.12.4debug30.txt/download](https://sourceforge.net/projects/maxbox5/files/examples/1417_statsmodels_64_delphi_python3.12.4debug30.txt/download)  
[https://sourceforge.net/projects/maxbox5/files/examples/1417\\_statsmodels\\_64\\_delphi\\_python3.12.4debug30.txt/download](https://sourceforge.net/projects/maxbox5/files/examples/1417_statsmodels_64_delphi_python3.12.4debug30.txt/download)



maXbox5 64-bit ScriptStudio 1417\_statsmodels\_64\_delphi\_python3.12.4debug30.txt

File Program Options View Debug Output Help

Load Find Replace / Refact Go Compile! Use Cases

```

386 //S.Delimiter := #10;
387 S.LineBreak := #10;
388 S1.Delimiter := ',';
389 //S.NameValueSeparator:= ',';
390 s.loadfromfile(exepath+'\examples\1417_export_dataframe.csv');
391 //S.DelimitedText:= s.text;
392 /// convert
393 TStringListToMatrix(s, amatx);
394 for i:= 0 to High(amatx)-80 do begin
395   for j:= 0 to High(amatx[i]) do
396     Write(amatx[i][j] + ' ');
397   Writeln('');
398 end;

```

maXbox5 C:\maxbox\maxbox5\examples\1417\_statsmodels\_64\_delphi\_python3.12.4debug30.txt Ct:01/07/2025 20:39:19 Mem:72% Row: 392-Col: 17 s:1401 M!

```

=====
Dep. Variable:    Lottery    R-squared:        0.414
Model:            OLS        Adj. R-squared:    0.392
Method:            Least Squares    F-statistic:      19.30
Date:            Tue, 01 Jul 2025    Prob (F-statistic): 1.47e-09
Time:            20:41:36        Log-Likelihood:    -375.28
No. Observations: 86          AIC:              758.6
Df Residuals:     82          BIC:              768.4
Df Model:         3
Covariance Type:  nonrobust
=====

```

	coef	std err	t	P> t	[0.025	0.975]
Intercept	194.3951	37.768	5.147	0.000	119.263	269.527
Wealth	0.2820	0.093	3.024	0.003	0.097	0.468
Literacy	-0.3840	0.127	-3.033	0.003	-0.636	-0.132
np.log(Pop1831)	-25.2363	6.047	-4.174	0.000	-37.265	-13.207

Python for Delphi at maXbox statsmodels console

Posted in [Conference](#), [data-science](#), [EKON](#), [Engineering](#), [math](#), [maXbox](#), [Python](#) Tagged [ai](#), [artificial-intelligence](#), [Coding](#), [data-science](#), [Machine Learning](#), [Python](#), [Statistics](#) [Leave a comment](#)

[Create a free website or blog at WordPress.com.](#)