International conference “Challenges and realities of the IT space: software engineering and cyber security”, 2022, October.
Online Scientific Publishers OWN, ISBN 978-83-7712-049-1.
Url:https://easychair.org/publications/preprint/7Cm9.

T. A. Chupilko, Yu. V. Ulianovska, M. F. Mormul, O. M. Shchitov

PYTHON FOR DATA PROCESSING AND MODELING INDICATORS OF ECONOMIC SECURITY OF THE COUNTRY

University of Customs and Finance, Dnipro

Abstract. The article considers aspects of efficient data processing. There are defined the stages of working with data and the features specific to each stage. There are considered packages NumPy, Pandas, Matplotlib, SciPy, Statsmodels and Scikit-learn. An example of using Python for customs tasks is given, taking into account the indicators of the country's economic security. The authors have created a calculation program that applies the above packages. A line of regression models are built to analyze the replenishment of the state budget of Ukraine with customs revenues at the expense of import and export duties. The analysis of models on the basis of econometric methods of modeling is carried out and forecast estimates of revenues are calculated.

Key words: Python, data processing, modeling, forecasting, regression model.

Introduction

In modern world, companies, enterprises, and institutions deal with a large amount of data. Recently, technologies that allow working with large amounts of information are gaining more and more popularity. The approach to data processing depends, first of all, on its type, purpose of use, capabilities of the enterprise or institution in the organization of data collection, systematization, analysis, which are often limited. Actually, a small number of enterprises and institutions are able to invest significant funds in the development of these technologies for their own needs. Undoubtedly, large companies have opportunities for the development of business analytics. For most techniques of data processing, the amount of data can be both large and small. How big data is needed for the result is determined by the interests of the company. Companies do not always deal with large volumes of generated data. Most often, there is a certain database at the company that must be used.

The main analysis techniques include cluster and factor analysis, modeling and forecasting based on econometric and optimization methods, outlier detection, artificial intelligence, network graphs, and machine learning. Some of these methods were developed for a long time, while others have appeared recently. At the right choice, various techniques and technologies are equally effective for evaluating the situation in various spheres of activity and making appropriate management decisions. Therefore, it is important to understand which technique is suitable for solving certain problem, how these techniques are used, and how they can be applied to model the indicators.

Similar basic mathematical tools are used in various technologies for studying data: both in standard data processing packages and corresponding libraries and modules in the popular object-oriented language Python, and in the widely known and widespread MS Office package, and in other software products such as the R language that is widely used for statistical data analysis.

In the pointed topic of the paper the authors used the works of foreign authors [1], [2], where the appication of Python and software packages specifically for data analysis are considered. In published articles [3] - [5] there are raised the issues related to the problems of this investigation. There were published several papers works by Ukrainian and foreign scientists regarding considerations about access to big data, their implementation in statistics, their real benefit, etc. Despite the significant scientific achivements on general issues about data, there is

a lack of researches with application of tools for effective data processing using programming languages, particularly, in the field of customs.

Topicality

Modeling covers various fields and different indicators. Especially important is the modeling of financial and economic indicators, which makes it possible to forecast their value in the presence of a certain trend, to provide a point and interval estimation of the prediction, which is determined by a confidence interval. Special criteria make it possible to assess the quality of the constructed model. Therefore, data analytics helps to improve the decision-making process in any field of activity. Але, разом з тим, потрібні інструменти для високоефективної та швидкої обробки даних. However, at the same time, highly efficient and fast data processing requires tools.

Recently, the Python programming language and a large number of open source libraries, which are dynamically updating, are a very popular and powerful tool that allows you to effectively process data, model and predict indicators using the interoperable ability to write code and ready-made solutions.

Data processing technologies are determined by the type of the data and the purpose of the research.

However, application of various technological tools have the common problems - data selection and preparation, as well as professional processing of the results.

In order to use even a completely automated data processing system, it need to be correctly selected, sorted, normalized, then an analysis method should be choosen, and after the data is processed by the program, it is necessary to carry out the analysis, interpretation, prediction, etc.

At the stage of preparing data, we are facing with many problems related to different data formats retrieved from different data sources, limited access to data because of the value of data, its confidentiality and strict regulation. Data can have different units of measurement as well as different levels of aggregation.

Improving the quality of data, understanding how data interact with each other, evaluating distributions and conversion them to a certain format is impossible without knowledge of the fundamentals of the corresponding mathematical tools.

Modern packages support many popular techniques of modeling and model evaluation. The application of the modern packages requires knowing of the corresponding programming language and mastering the packages, in which you need to set various parameters to use a certain method or function..

Another problem is related to data access. Official statistical data, which are the basis for modeling and predicting financial and economic indicators on a national scale, in particular in the customs area, are quite limited and, mainly, aggregated. Often, official websites or statistical collections provide data that are not normalized, or their quantity is insufficient to build adequate models. Therefore, modeling tasks are limited to those statistical data that are officially available.

The modeling process depends on the quality of the data, as well as on the professionalism of the analyst.

Aim

The purpose of the work is to analyze effective tools for data processing and modeling; application of Python and libraries for analysis, modeling and predicting of financial and economic indicators using the example of official data of customs revenue to the state budget of Ukraine on export and import duties.

Tasks

1. Analyze problems in working with data and aspects of application of the tools for effective data processing.
2. Apply the most applicable tools for the specific task of modeling financial and economic indicators (for example, customs revenues to the budget from individual indicators).

Solving problems

Tools for efficient and fast data processing

For data processing, it is appropriate to use techniques that are most effective for a certain type of the solving problem.

In some tasks, it is sufficient to use a convenient, understandable and accessible tool, such as MS Excel that has a line of capabilities and includes an analysis package that despite of its limitation nevertheless is suitable for obtaining results of data processing in the first approximation to present the nature of the data. Using spreadsheets, it is impossible to launch a production model, for example, artificial intelligence, in software mode. However, spreadsheets allow you to analyze the nature of the data, simulate and predict the results. The result can be obtained on the basis of classical approaches of probability theory and mathematical statistics for data normalization, correlation and regression analysis, finding the point and interval estimations of the values, as well as using procedures for determining optimal solutions of linear and non-linear optimization problems.

Actually, the skills in one or more programming languages are required to use the data processing automation in a program way. However, it is not necessary to know how the code is written in order to understand the essence of the analysis used in processing technologies and various application packages such as Statistics, SPSS, etc. These powerful tools include a variety of analyses, including regression analyses, factorial analyses, cluster analyses, model building by means of neural networks, and much more, as well as give an opportunity to get graphical representation of the results in case the dimension and formulation of the problem allow it.

In the process of working with data, several stages can be highlighted.

1. Purpose of the aim of the study. At the same time, a project task is prepared as well as the purpose of the research and the cost of the work are evaluated.
2. Collection and preparation of data, the so-called "intelligence analysis". Virtually, certain difficulties arise already at this stage. Data can be disparate, in different formats and need to be normalized and brought to uniformity. Matrices may be incompletely filled and singular. It is necessary to choose an algorithm for filling the gaps. Often there are significant deviations in the data, that is, outliers that need to be eliminated, that is, the data must be cleaned. Otherwise, no modelling methods will lead to an adequate model. Therefore, the data preparation process is very painstaking and routine, almost "manual", it requires an intellectual approach and an understanding of the purpose of the study.
3. Analysis and data modelling, namely selection of a model and evaluation of its parameters, in machine learning - so-called "model training". At this stage, it is necessary to understand how the data are related to each other, estimate the distributions of the data, identify and eliminate outliers, as well as check for multicollinearity in the system and for negative phenomena such as heteroskedasticity and autocorrelation, which require additional transformations of variables and application of the special methods. Certain statistical methods and simple modeling are used for this purpose. However, a number of questions arise: are the studied factors and the indicator related to each other, is there multicollinearity in the data system, is it possible to reduce the number of variables and thereby simplify the model, what form of dependence should be chosen for modeling, how to reduce the model

to a linear form, etc. At this stage, it is required the knowledge of the subject area, as well as experience in mathematics, probability theory and mathematical statistics. Only after carrying out the specified studies and data transformations it is possible to use ready-made solutions - program packages. The modeling process itself, called "model training", means building different models on one set of data, randomly selected from the general population. The number of data can be varied using the parameters set for the selected method. It can be choosen the best model according to certain criteria, for example, the method of least squares, or a method based on a decision tree with different parameters that can be varied, or the method of absolute deviations, etc.

The data set can be trained several times, changing the parameters, and thus achieving the best results. Therefore, building a model is an iterative process and requires the skills of a researcher.

4. Checking the adequacy of the model and the significance of the factors of the model. After obtaining the best result (for example, the sum of squared residuals is comparing and the set of parameters that gives the smallest sum is selecting), the quality of the model is evaluated according to statistical criteria. If the quality is unsatisfactory, then the model needs to be retrained.
5. Application of the model to unknown data - the so-called "training set" is selected from the same sample - "predictive modeling", that is, the forecast is done.

The described approach is used for modeling and forecasting tasks in machine learning. Python, for example, has its own Scikit-learn library with a variety of algorithms.

Novadays, machine learning is a very popular and promising technology among analysts (data-scientists). The machine learning market is growing rapidly. Since 2016, its volume has exceeded $1 billion, and according to forecasts, it can grow to $39.98 billion by 2025. 60% of companies in the world already use machine learning.

Among the tasks that can be solved by means of machine learning, we can mention the tasks of modeling and forecasting of the indicators depending on one or more factors as well as optimization tasks. There are used both traditional methods of econometric analysis, including single-factor and multi-factor models based on the method of least squares, and non-traditional ones, such as decision trees with a large number of adjustable parameters, which provide flexibility in modeling model parameters.

The so-called "neural networks" are getting more and more popular. In modeling, the concept of risk is used, the quantitative characteristics of which are calculated according to the numerical characteristics of discrete and continuous random variables.

Over the past ten years, Python has become one of the most important programming languages used in data science, machine learning, and general-purpose software development in academic institutions and industry. Improved libraries for Python led to the fact that it became a serious competitor in solving the problems of creating data processing applications.

In many modern environments, it is used a common set of legacy libraries written in FORTRAN and C, that contain implementations of algorithms for linear algebra, optimization, integration, etc. Therefore, many companies use Python as "glue" to combine programs written over many years.

Python packages for working with data

NumPy, abbreviation from "Numerical Python", is a basic package for performing scientific calculations in Python. Other libraries are built on the base of NumPy.

The main features of the package are the following: it can be created quickly and efficiently the objects of multidimensional arrays ndarray ; it has functions for performing calculations with elements of one array or mathematical operations with several arrays; it provides ways for reading and writing to disks data sets presented in the form of arrays; it uses

linear algebra operations, Fourier transform and random number generator; it has facilities for integration with code written in C, C++, or Fortran.

NumPy significantly accelerate work with arrays. As a means of data storing and data manipulating NumPy arrays are much more efficient than Python's built-in data structures.

Many Python-oriented computing tools either use NumPy arrays as the underlying data structure, or somehow else arrange integration with NumPy.

Pandas provides data structures and functions that make working with structured data more simple and fast. Due to to this library, Python has turned into a powerful and productive data analysis environment. The main pandas objects are *DataFrame* - a two-dimensional table in which rows and columns have labels, and *Series* - a one-dimensional array object with labels.

The pandas library combines the high performance of NumPy's array tools with the flexible data manipulation capabilities of spreadsheets and relational databases (eg, SQL-based). Since data manipulation, preparation and cleaning have a very big role in data analysis, pandas is one of the main tools.

The main capabilities of the library are the following: it has advanced indexing tools that allow you to easily change the shape of data sets, form slices, perform aggregation and select subsets; data structures with marked axes support automatic or explicit data alignment, which eliminates typical errors when working with unaligned data and data from different sources that are indexed differently; it has built-in time series functionality; the same data structures can support both time series and other types of data; arithmetic operations, which are performed with objects as with numerical data; it has flexible processing of missing data (supplementation); data integration; support for connection and other relational operations available in popular databases (for example, based on SQL).

Many facilities present in pandas are either part of the R language or provided by additional packages.

The title pandas itself is derived from panel data, used in econometrics to denote multidimensional structured data sets, and from the phrase Python data analysis.

Matplotlib is the most popular tool in Python for creating graphs and other ways of visualizing two-dimensional data, suitable for creating graphs convenient for publication. Although there are visualization capabilities in other packages, matplotlib is used most often and is therefore well integrated with other parts of the ecosystem.

SciPy is a set of packages intended to solve various standard computing problems. Some of them: scipy.integrate - routines for numerical integration and solving differential equations; scipy.linalg - subroutines for linear algebra and matrix expansion, supplementing those included in numpy.linalg; scipy.optimize - algorithms for optimizing functions (finding extrema) and finding roots; scipy.signal - signal processing tools; scipy.sparse - algorithms for working with sparse matrices and solving sparse systems of linear equations; scipy.special - a wrapper around SPECFUN, written in the Fortran library, containing implementations of many standard mathematical functions, including the gamma function; scipy.stats - standard continuous and discrete probability distributions (probability density functions, sample formation, continuous probability distribution functions), various statistical criteria and additional descriptive statistics.

Scikit-learn is the core Python machine learning toolkit for programmers. It has submodules for the following models: classification: support vector method, nearest neighbor algorithm, random forests, logistic regression, etc.; regression: Lasso, ridge regression, etc.; clustering: k-means method, spectral clustering, etc.; dimensionality reduction: method of principal components, feature selection, matrix factorization, etc.; model selection: grid search, cross-checking, metrics; preliminary processing: feature selection, normalization.

Scikit-learn is mainly focused on forecasting and prediction.

Statsmodels is statistical analysis package. Compared to Scikit-learn, the Statsmodels package contains algorithms for classical (primarily frequency) statistics and econometrics. It includes the following submodules: regression models: linear regression, generalized linear models, linear models with mixed effects, etc.; analysis of variance (ANOVA); time series analysis: AR, ARMA, ARIMA, VAR and other models; non-parametric methods: kernel density estimation, kernel regression; visualization of statistical modeling results.

The statsmodels package is focused more on statistical inference, it provides uncertainty estimates and p-values of parameters. It used along with with NumPy and Pandas.

Python has libraries for convenient and quick data reading in the formats of spreadsheets, databases, csv, etc.

An example of application of Python to model customs revenues to the state budget of Ukraine

For this research there were used data from official statistics [6].

It should be noted that the information in public access is very limited.

There are consolidated data for the task, which include the full volume of receipts from customs authorities to the state budget of Ukraine, as well as receipts from import and export duties, shown in the table 1.

Table 1 – raw data for modeling receipts to the state budget of Ukraine from customs authorities in total and by individual types.

Year	Income to the state budget of Ukraine Y, UAH	Income to the state budget of Ukraine by import duty X1, UAH	Income to the state budget of Ukraine by export duty, X2, UAH
2013	1,40036E+11	1,2550E+10	2,4900E+08
2014	3,57084E+11	1,3056E+10	1,7900E+08
2015	5,34694E+11	1,7422E+10	2,4500E+08
2016	6,16219E+11	2,0004E+10	3,7000E+08
2017	6,98405E+11	2,2257E+10	6,4300E+08
2018	8,33615E+11	2,3301E+10	5,1600E+08
2019	8,79833E+11	2,2778E+10	2,3000E+08
2020	8,77603E+11	2,1538E+10	2,5700E+08

We will conduct an analysis of individual components of incomes, in particular, import and export duties.

According to the data in Table 1, we will evaluate the presence and closeness of the relationships between them and the total incomes from the customs authorities, the form and type of the model, the regression parameters, the adequacy of the model, the statistical significance of the parameters, the presence of autocorrelation, we will determine the predictive value of the indicator (point and interval estimation) and construct regression confidence intervals. We will apply Python and NumPy, Statsmodels, Matplotlib, Xlrd libraries as a modelling tool (for reading data from an Excel file).

The most common modern method of modeling structured data is econometric modeling. It is estimated the dependence of the indicator on one or more factors using the regression analysis. The best result is given by the method of least squares of deviations of the original

data from the simulated ones. The adequacy of the model will be assessed using the Fisher test. We will evaluate the statistical significance of the regression parameters, as well as the confidence intervals of the regressions, based on the Student's test. In addition, we will obtain other statistics based on the model and apply the model for prediction.

Below are listings of the results of software execution of calculations in Python and graphics for a visual representation of the constructed models.

Initially, let's determine how the incomes to the budget from the customs authorities depend on the income for the import duty. Fig. 1 and Fig. 2 show the performance results.

Let's analyze the main simulation results.

The OLS model, as well as the Least squares method, was applied. The dependent variable is y. The regression equation is shown in Fig. 1. The correlation coefficient is 0.95 indicating a strong correlation between the factor and the indicator. The adjusted coefficient of determination is 0.885: the change in the indicator is due to the change in the factor by 88.5%.

The F-statistics show the adequacy of the model: the calculated value is 55.03, the critical value for the degrees of freedom of the problem and the significance level of 0.05 is 5.98. The calculated t-statistic values are 7.42 for the slope and -3.25 for the regression intercept. Both parameters are statistically significant with a confidence probability of 0.975. The critical value of the t-statistic is 2.45.

Thereafter, the confidence intervals of the regression parameters at a significance of 0.025: for the slope: (39.14; 77.67), for the intercept :(-8.75E+11; -1.23E+11). The Durbin–Watson statistic indicates the absence of autocorrelation in the model. The covariance matrix is correctly specified.

The model can be used to predict the indicator. Forecast estimates (point and interval) are defined. The coefficient of elasticity according to the average indicators for the last four years is equal to 1.69 meaning that the indicator is elastic by factor, and the rate of growth of filling the budget from customs revenues is slowing down, in particular, due to import duties.

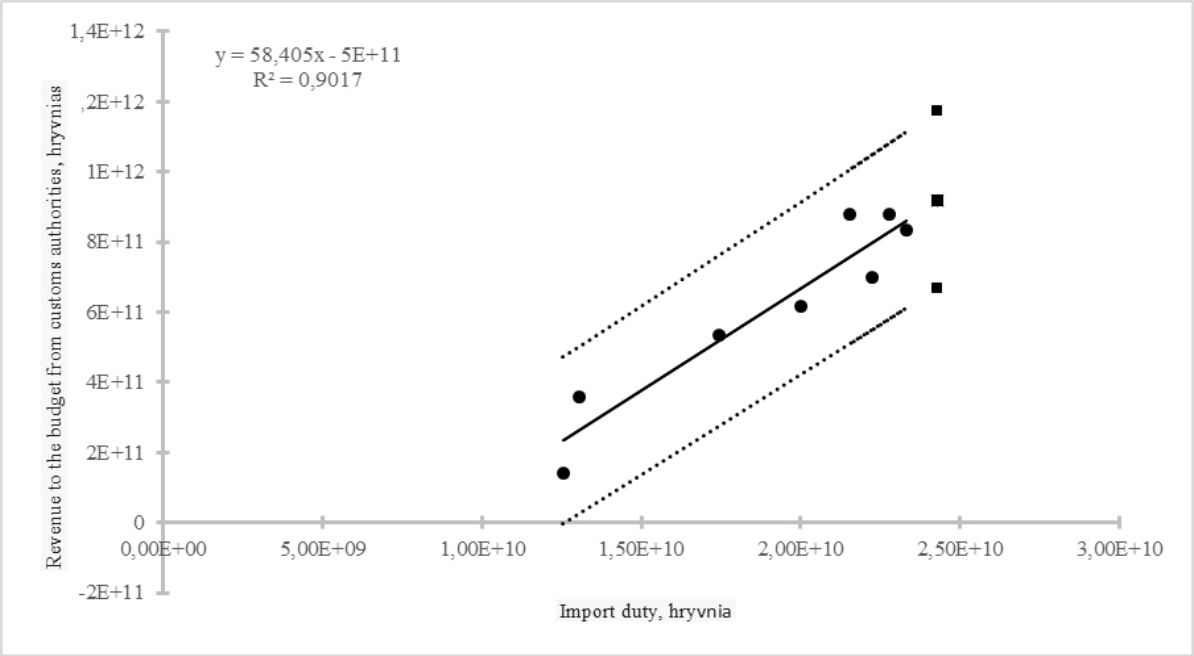

Figure 1 – Regression of incomes to the budget of Ukraine from customs authorities on import duty, confidence interval of the forecast, confidence interval of the regression, constructed with a reliability of 0.95

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.902
Model:                            OLS   Adj. R-squared:                  0.885
Method:                 Least Squares   F-statistic:                     55.03
Date:                Tue, 03 Aug 2021   Prob (F-statistic):           0.000309
Time:                        22:34:43   Log-Likelihood:                -211.97
No. Observations:                   8   AIC:                             427.9
Df Residuals:                       6   BIC:                             428.1
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
x1            58.4053      7.873      7.418      0.000      39.141      77.670
const      -4.991e+11   1.54e+11     -3.245      0.018   -8.75e+11   -1.23e+11
==============================================================================
Omnibus:                        1.280   Durbin-Watson:                   1.337
Prob(Omnibus):                  0.527   Jarque-Bera (JB):                0.653
Skew:                           0.145   Prob(JB):                        0.722
Kurtosis:                       1.631   Cond. No.                     9.46e+10
==============================================================================

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified
[2] The condition number is large, 9.46e+10. This might indicate that there are
strong multicollinearity or other numerical problems.
Correlation
[[1.     0.9496]
 [0.9496 1.    ]]
Factor value for a given indicator forecast [[2.56681055e+10]
 [2.85201172e+10]]

Process finished with exit code 0
```

Figure 2 – List of execution of the program (model of dependence of customs revenue to the budget on the import duty)

We will construct and analyze the regression of customs revenues to the export duty budget. Similarly to what was explained earlier, we will get the result of the program execution. For a visual presentation of the result, we have a diagram with a regression equation in Fig. 3.

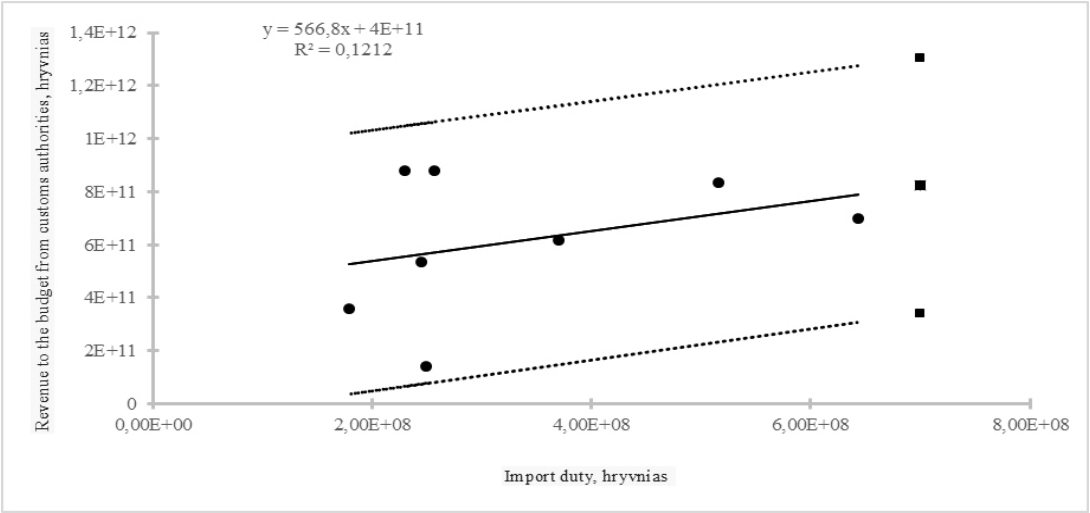

Figure 3 – Regression of incomes to the budget of Ukraine from customs authorities on export duty, confidence interval of the forecast, confidence interval of the regression, constructed with a reliability of 0.95

Figure 4 shows the corresponding listing.

```
Результат розрахунку:
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.121
Model:                            OLS   Adj. R-squared:                 -0.025
Method:                 Least Squares   F-statistic:                    0.8276
Date:                Tue, 03 Aug 2021   Prob (F-statistic):              0.398
Time:                        22:29:22   Log-Likelihood:                -220.73
No. Observations:                   8   AIC:                             445.5
Df Residuals:                       6   BIC:                             445.6
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
x1           566.7962    623.032      0.910      0.398    -957.708    2091.300
const       4.267e+11    2.3e+11      1.856      0.113   -1.36e+11    9.89e+11
==============================================================================
Omnibus:                        0.138   Durbin-Watson:                   0.410
Prob(Omnibus):                  0.933   Jarque-Bera (JB):                0.213
Skew:                          -0.204   Prob(JB):                        0.899
Kurtosis:                       2.313   Cond. No.                     8.94e+08
==============================================================================

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified
[2] The condition number is large, 8.94e+08. This might indicate that there are
strong multicollinearity or other numerical problems.
Correlation
[1.     0.3482]
 [0.3482 1.    ]]
Factor value for a given indicator forecast                : [[4.44254957e+08]
 [4.95616619e+08]]

Process finished with exit code 0
```

Figure 4 – List of program implementation (model of dependence of customs incomes to the budget on export duty)

The same OLS model as in the previous case is applied. The dependent variable is y. Рівняння регресії виведено на рис.3. The correlation coefficient is equal to 0.348 that indicates a very weak correlation between the factor and the indicator. The coefficient of determination of 0.1212 is very low, close to zero: the change in customs revenues is caused by a 12% change in export customs revenues. The F-statistics indicates the inadequacy of the model: the calculated value is 0.83, the critical value for the degrees of freedom of the problem and the significance level of 0.05 is 5.98. Calculated t-statistic values are 0.91 for the slope and 1.86 for the regression intercept. Both parameters are statistically insignificantly different from zero with a significance level of 0.125. The critical value of the t-statistic is 2.45 for the degrees of freedom of the model and the given level of significance.

The very wide confidence interval is due to the large deviation of the original data, and accordingly, the wide confidence interval of the prediction is of no practical value. Durbin–Watson statistics indicate the absence of autocorrelation in the model. The covariance matrix is correctly specified.

The model, obviously, cannot be recommended to be used to the prediction of the indicator.

The coefficient of elasticity according to the average indicators for the last four years is equal to 0.34, increasing to the value of 0.5, that is, the indicator is inelastic in terms of the factor, and the growth rate of filling the budget from customs revenues is accelerating due to the export duty.

Let's draw general conclusions about the last two models. For both models, the least squares method was used which gives the best approximation of the original data with the least sum of squared errors. Estimates obtained by this method according to the Gauss-Markov theorem are efficient and unbiased. For both models, the calculation was done applying the Python program created by the authors and using NumPy, Statsmodels, Matplotlib, and Xlrd libraries. Preliminary data preparation was done in Excel.

It is shown in the article the possibility of software processing of data. It was not possible to obtain more complete statistical data from the official website to be able to conduct an analysis on large arrays of similar data. But the problems that appeared during modelling based on the given aggregated data would be reproduced similarly. In addition, the listings contain a warning that the amount of data for a correct calculation must be larger. If the amount of data is insufficient, then it is necessary to use adjusted estimates, which is also known from the theory of mathematical statistics.

It can be noted that with a small amount of data the statistical analysis provided by Excel could be used, and it would be the most effective solution. But the purpose of the current work was to apply the capabilities of Python for modelling. The same program would give an efficient calculation for large arrays of data, which is not a problem in programming directly. It can also be noted that data can be stored in databases. Python has appropriate tools for working with such data. Therefore, we have relevant economic conclusions, which reveal big problems in incomes from customs authorities and, in particular, from export duties.

It is possible to build various single-factor and multi-factor: linear and non-linear models, and in this way, get a complete picture of what processes are taking place in the industry, in particular, in customs, and identify positive and negative phenomena. Based on the results of modeling, it is possible to obtain scientifically based prediction and make appropriate management decisions.

Conclusions

1. An analysis of problems arising in data processing and effective tools for data modelling and prediction has been carried out.
2. There were built the models for the analysis of customs revenues to the state budget of Ukraine at the expense of import and export duties. The Python programming language and packages appropriate to the tasks were used.
3. The possibility of using the results of modeling financial and economic indicators for making management decisions is indicated.

References

[1] U. Makkyny, *Python y analiz dannykh.* M., Rossia: DMK Press, 2020, 540 s.

[2] S. Devy, M. Arno, A. Mokhamed*, Osnovy Data Science i BigData. Python y nauka o dannykh*. Peterburg, Rossia: Pyter, 2017, 336 s.

[3] T.A. Chupilko, "Aktualni problemy vysokoefektyvnoi obrobky danykh. Modeliuvannia pokaznykiv za dopomohoiu movy prohramuvannia Python" u *Aktualni napriamy rozvytku tekhnichnoho ta vyrob-nychoho potentsialu natsionalnoi ekonomiky.* Dnipro: Porohy, 2021, s.151-163

[4] T.A.Chupilko, "Bazovyi instrumentarii u suchasnykh tekhnolohiiakh kompiuternoi biznes-analityky" u *Mizhnar. Nauk. Konf. Innovatsiini tekhnolohii, modeli uprav-linnia kiberbezpekoiu ITMK-2020*, Dnipro, 2020, T2, s. 53-54

[5] T.A. Chupilko, "Kompiuterni tekhnolohii ta ekonomiko-matematychni metody v upravlinni biznes-protsesamy na pidpryiemstvi" u *Mizhnar. Nauk. Konf. Innovatsiini tekhnolohii, modeli upravlinnia kiberbezpekoiu ITMK-2020*, Dnipro, 2020, T.1, s. 26-28

[6] Ministerstvo finansiv Ukrainu. [Electronic resource]. Access mode: http://mof.gov.ua. Data of application: 20 serpnya, 2021.

СЕКЦІЯ: БЕЗПЕКА ТА РОЗРОБКА ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ