

**ՏՎՅԱԼՆԵՐԻ ՎԵՐԼՈՒԾՈՒԹՅՈՒՆԸ ՏՆՏԵՍԱԳԻՏՈՒԹՅԱՆ ՄԵՋ PHYTON  
ՎԻՃԱԿԱԳՐԱԿԱՆ ՎԵՐԼՈՒԾԱԿԱՆ ԾՐԱԳՐԱՅԻՆ ՓԱԹԵԹԻ  
ՕԳՏԱԳՈՐԾՄԱՄԲ**

**Գևորգ ԳՐԻԳՈՐՅԱՆ**

ԿԱԻ գլխավոր հաշվապահ, տ. գ. թ., դոցենտ

**Ահարոն ՌՈՒՇԱՆՅԱՆ**

ՀԱՊՀ ՏՀՏԷԻ, ՏԱԾԱ ամբիոնի, Ծր. ճարտ. մագիստրոս)

Բանալի բառեր. տնտեսագիտական վերլուծություն, Python, տվյալների վորլուծություն, Pandas, տնտեսագիտական կանխատեսում, տնտեսագիտական մոդել

Տնտեսագետի գործունեության առանձնահատկությունն այն է, որ աշխատում է տվյալների մոդելավորման համար նախնական տվյալների բացակայության պայմաններում: Նման տեղեկատվության վերլուծությունը պահանջում է հատուկ մեթոդների կիրառում, որոնք կազմում են տնտեսագիտության ասպեկտներից մեկը: Սովորաբար, տնտեսագետներն աշխատում են ավելի քիչ հաճախականության տվյալների հետ և օգտագործում են ավելի շատ ընդունված տեսական հիմքեր, մինչդեռ տվյալների մշակման մասնագետները աշխատում են բարձր հաճախականության տվյալների հետ և ցիկլային տրվյալների թեստերով [1] [Բեկոն 2018թ. էջ 1]: Տնտեսագիտության կարևոր խնդիրներից է տնտեսաչափական մոդելի կառուցումը և մոդելի օգտագործման հնարավորության որոշումը՝ տնտեսական գործընթացների նկարագրման, վերլուծության և կանխատեսման համար [2] [Կասպարովիչ 2007թ. էջ 3]:

Կառուցվածքային և չկառուցված տվյալների կառավարումը, դրանց մշակման համար նոր տեխնոլոգիաների և գործիքների օգտագործումը մի թեմա է, որը ներկայումս արդիական է ինչպես կոնկրետ ձեռնարկությունների, այնպես էլ պետական մակարդակով: Ըստ International Data Corporation-ի որը հանդիսանում է միջազգային հետազոտական և խորհրդատվական ընկերություն՝ կնխատեսվում է մեծ տվյալների մշակման տեխնոլոգիաների և բիզնես-անալիտիկայի շուկայի տարեկան աճ 11,9%-ով:[3] [Tadviser ընկերություն 2020թ]: Սկզբի համար սպասվում է Big Data վերլուծական հարթակների աճ՝ որոնք օգտագործում են մեքենայական ուսուցման և արհեստական ինտելեկտի համակարգերը՝ հիմնված նյարդային համակարգերի վրա, որոնք

պահանջված են բիզնեսի կողմից՝ սպառողների մասին տվյալները վերլուծելու և նրանց վարքագիծը կանխատեսելու համար:

Վերջին շրջանում բազմաթիվ հետազոտություններ են իրականացվել, այսպես կոչված, մեծ սոցիալական տվյալների ոլորտում: Մի շարք փորձագետներ այս ոլորտը դիտարկում են որպես առանձին գիտություն: Այստեղ կարելի է առանձնացնել չորս նշանակալից ճյուղեր՝ սոցիալական հաշվարկ (social computing), մեծ տվյալների մասին գիտություն (Big Data science), տրվյալների վերլուծություն (data analytics) և հաշվողական սոցիոլոգիա (computational social science, CSS): Տվյալների վերլուծությունը զբաղվում է ամենաընդհանուր մաթեմատիկական մեթոդների և հաշվողական ալգորիթմների կառուցմամբ և հետազոտմամբ՝ փորձարարական (լայն իմաստով) տվյալներից գիտելիք հանելու համար, տվյալների հետազոտման, զտման, փոխակերպման և մոդելավորման գործընթաց է՝ օգտակար տեղեկատվություն ստանալու և որոշումներ կայացնելու համար: Տվյալների վերլուծության տեխնոլոգիաների օգտագործումը մանրաձախ առևտրի և սպասարկման ոլորտի խոշոր և միջին ընկերությունների համար մեծ առաջընթաց է: Այս տեխնոլոգիաները ակտիվորեն օգտագործվում են բանկերի և բջջային օպերատորների կողմից: Բացի այդ, դրանք օգտագործվում են խոշոր արտադրական ընկերությունների կողմից՝ վերլուծելու սարքավորումների խափանումների տվյալները և կլրճատելու պարապուրդը, որն օգնում է նվազեցնել ծախսերը: Օրինակ, թռիչքների կառավարման ոլորտում տվյալների հավաքածուների վերլուծությունը կարող է բարձրացնել սարքավորումների հուսալիությունը և նվազեցնել խափանումների թիվը: Տվյալների հավաքագրման և փորձարկման խնդիրները միշտ էլ մարտահրավեր է եղել տնտեսագետների համար, բայց մեր օրերում ավելի ու ավելի են հեշտացել նման խնդիրների լուծումը քանի որ կան բազմաթիվ ծրագրավորման լեզուներ, ինչպիսին է Python-ը որը բարելավվում է արդյունավետությունը և թույլ է տալիս վերլուծել ավելի բարդ մոդելներ:

Python-ը դա ընդհանուր նշանակության բարձր մակարդակի ծրագրավորման լեզու է, որը հիմնականում կենտրոնացած է ծրագիրը արագ մշակելու և կոդի հեշտ ընթերցանությունն ապահովելու վրա [4] [Python Software Foundation 2021թ. էջ 1]: Python-ի միջուկի շարահյուսությունը շատ պարզ և հեշտ է: Միևնույն ժամանակ, ստանդարտ գրադարանն ներառում է բազմաթիվ օգտակար ֆունկցիաներ: Python-ն աջակցում է ծրագրավորման մի քանի

նմուշների, այդ թվում՝ կառուցվածքային, ֆունկցիոնալ, իմպերատիվ և ասպեկտային, օբյեկտային կողմնորոշմամբ: Հիմնական կառուցվածքային գրծերն են՝ դինամիկ տիպավորումը, հիշողության ավտոմատ կառավարումը, բացառությունների վերամշակման մեխանիզմը և բարձր մակարդակի տվյալների կառուցվածքը: Python-ում կողը կազմված է ֆունկցիաներից և դասերից, որոնք կարող են միավորվել մոդուլներում, որոնք էլ, իրենց հերթին, փաթեթներում: Այն օժանդակում է օբյեկտային ուղղվածությամբ ծրագրավորմանը և վերջերս դարձել է ծրագրավորման լեզուների լուրջ այլընտրանք օրինակ՝ C ++ -ին [5] [Pandas development team 2021թ. էջ 5]: Pandas-ը համարվում է Python ծրագրավորման լեզվի գրադարան: Գրադարանը թույլ է տալիս հեշտությամբ աշխատել նախնական տվյալների հետ և դրանք վերլուծել: Ըստ էության, Pandas-ի գրադարանը կառուցվում է մեկ այլ հիանալի գրադարանում Python 3 - Numpy-ում: Pandas փաթեթի օգտագործումը լայն հնարավորություններ է տալիս աղյուսակների հետ աշխատելու համար:

Տվյալների վերլուծության հետ աշխատելու համար անհրաժեշտ է ֆունկցիոնալ և միևնույն ժամանակ պարզ լեզու: Պարզության և բարձր ֆունկցիոնալության հետ մեկտեղ, Python-ն ունի շատ այլ առավելություններ՝

- Python սովորելը դժվար չէ և շատ ավելի քիչ ժամանակ կպահանջի, քան Java, C և այլն սովորելը,

- պարզ գրելաձևը թույլ է տալիս արագ գրել կողը: Շատ ավելի հեշտ է դա անել Python-ում, քան Java-ում կամ C-ում: Այո, Python-ում գրված կողը հեշտ է կարդալ և մեկնաբանել,

- Python-ն ունի ներկառուցված թարգմանիչ, որը թույլ է տալիս արագ գրել կողը: Սա վերլուծաբանների համար շատ ավելի հեշտ է դարձնում բազմաթիվ վարկածների փորձարկումը, քանի որ թույլ է տալիս դա անել ինտերակտիվ կերպով: Այլ լեզուներով դրան հասնելը շատ դժվար է,

- Big Data-ի հետ աշխատող մասնագետներին անհրաժեշտ կլինի նաև ներկառուցված թարգմանիչ: Այն թույլ է տալիս վերափոխել աղբյուրի կողը մեքենայական հրահանգների և առաջացնել օպտիմալացման գաղափարներ,

- լեզվի զարգացման բարձր արագություն:

Համոզվելու համար Python-ը ինչպես կարող է բարելավել տվյալների վերլուծությունը տնտեսագիտության մեջ, կատարենք հետազոտություն: Դրա համար անհրաժեշտ է ընտրել հիմնական տվյալների հավաքածու՝ դա կլինի

DataSet-ը: Մի շարք տվյալների հավաքածուներ կարելի է ներբեռնել ուղղակիորեն այնպիսի կայքից, որը պարունակում է այդպիսի տեղեկատվություն: DataSet-ը, որպես կանոն, json կամ csv ֆորմատով ֆայլեր են, որոնք իրենցից ներկայացնում են աղյուսակով տվյալներ [6] [ RUVDS ընկերություն, 2020]:

Օրինակի միջոցով ներկայացնենք Python գործիքների օգտագործմամբ մեծ քանակությամբ տվյալների մշակման պարզությունը:

Օրինակի համար բեռնենք 2 տեքստային ֆայլ: Դա կարելի է անել read\_csv () գործառնությամբ:

```
from pandas import read_csv
df1 = read_csv("df1.txt") df2 = read_csv("df2.txt",";")
```

**Նկար 1** read\_csv () գործառնությամբ ֆայլերի ներբեռնում

Այժմ մենք ունենք 2 տվյալների հավաքածու df1, օրինակի համար գրենք տվյալներ, որոնք պարունակում են խանութների և առաքումների քանակը՝

shop	qty
427	3
707	4
957	2
437	1

**Նկար 2** Խանութների և առաքումների քանակաի աղյուսակ  
Եվ df2- ը, որը պարունակում է խանութը և նրա քաղաքը՝

shop	name
347	Киев
427	Самара
707	Минск
957	Иркутск
437	Москва

**Նկար 3** Խանութների և քաղաքների տվյալների աղյուսակ

Տվյալների հետ ծրագրի միջոցով կարելի է իրականացնել տարբեր գործառնություններ, ինչպիսիք են՝ միաձուլումը, սյունները ավելացնելը, գրառումները ավելացնելը, ֆիլտրելը, համախմբված սյունների կառուցում և այլն: Վերը նկարագրված գործառնությունները իրականացնելու համար հարկավոր է նշված օրինակում լուծել հետևյալ խնդիրները՝

Քաղաքների խանութների հավաքածուների մեջ ավելացնենք "country" դաշտը և լրացնենք համապատասխան երկրների անուններ: Ընտրենք օրինակ հայկական խանութ և փոխենք դրա քանակը: Նախորդ փուլում ձեռք բերված խանութը ավելացնենք ընդհանուր ցուցակին:

df1 քանակից ավելացնում ենք df2-ի քանակը, այնուհետև կառուցենք առանցքային սեղան ըստ երկրների և առաքումների քանակի:

Այսպիսով, կա insert () հրաման՝ նոր սյունակ տվյալների հավաքածուին ավելացնելու համար: Մեր դեպքում ֆունկցիան փոխանցվում է 3 արգումենտ՝

1. դիրքի համարը, որտեղ կտեղադրվի նոր սյունակը
2. նոր սյունակի անուն
3. սյունակների արժեքների զանգված (կանոնավոր ցուցակ)

Տվյալների հավաքածուն ֆիլտրելու համար կարող եք օգտագործել քառակուսի փակագծեր, որոնց ներսում կլինի տվյալների հավաքածուի [պայման] պայմանը: Պայմանը պետք է պարունակի դաշտերի անունները՝ վերը նկարագրված ձևաչափով և դրանց վրա դրված պայմանը: Այսպիսով, կարող ենք ընտրել, վերցնել խանութը և փոխարինել դրա համարը: Տվյալ կոդի կատարման արդյունքը կլինի նոր միջանկյալ տվյալների հավաքածու ստեղծումը, որը պարունակում է մեկ գրառում: Այսպիսով, կատարեցինք մեծ քանակությամբ տվյալների վերլուծություն և ֆիլտրացիա: Մեծ քանակությամբ տվյալների հետ աշխատելիս Python-ը Excel-ի լավ այլընտրանք է: Տվյալների վերլուծության համար որպես օրինակ օգտագործենք «Apple» ընկերության բաժնետոմսերի գինը 5 տարիների ընթացքում օրվա կտրվածքով:

```
>>> import pandas as pd
>>> df = pd.read_csv('apple.csv', index_col='Date', parse_dates=True)
>>> df = df.sort_index()
>>> print(df.info())
<class 'pandas.core.frame.DataFrame'>
DatetimeIndex: 1258 entries, 2017-02-22 to 2012-02-23
Data columns (total 6 columns):
Open          1258 non-null float64
High          1258 non-null float64
Low           1258 non-null float64
Close         1258 non-null float64
Volume        1258 non-null int64
Adj Close     1258 non-null float64
dtypes: float64(5), int64(1)
memory usage: 68.8 KB
```

**Նկար 4** DataFrame-ի ստեղծում DatetimeIndex ու

Այստեղ մենք ստեղծում ենք DataFrame-ը DatetimeIndex-ով Date սյունակում և տեսակավորում ենք նոր ինդեքս ճիշտ հերթականությամբ: Եթե սյունն ունի ամսաթվի և ժամանակի համապատասխան ձևաչափ ապա pandas.to\_datetime մեթոդը կարող է օգտագործվել տողը ճիշտ ձևափոխելու համար ցանկալի տիպի: Պարզենք բաժնետոմսերի միջին գինը (mean) փակման ժամանակ (close):

```
>>> df.loc['2012-Feb', 'Close'].mean()
528.4820021999999
```

**Նկար 5** Փակման ժամանակ բաժնետոմսերի գինը

Եթե մենք վերցնենք 2012թ. փետրվարից մինչև 2015թ. փետրվար ժամանակահատվածը և հաշվարկենք միջինը, կստացվի հետևյալ արդյունքը՝

```
>>> df.loc['2012-Feb':'2015-Feb', 'Close'].mean()
430.43968317018414
```

**Նկար 6** Բաժնետոմսերի միջին գնի հաշվարկ Շաբաթվա ընթացքի միջին գինը՝

```
>>> df.resample('W')['Close'].mean()
Date
2012-02-26    519.399979
2012-03-04    538.652008
2012-03-11    536.254004
2012-03-18    576.161993
```

**Նկար 6** Շաբաթվա ընթացքի միջին գնի հաշվարկ

Այսպիսով Python-ը արդյունավետ և հեշտ գործիք է ժամանակային տվյալների վերլուծության համար [7 Կուզնեցով 2013թ. էջ 3]: Տվյալների վերլուծությունը կիրառվում է տնտեսության գրեթե բոլոր ոլորտներում՝ ֆինանսական և SS ոլորտից սկսած մինչև, նավթի և գազի արդյունաբերության ձեռնարկություններ: Տեղեկատվության ծավալն անընդհատ աճում է ամբողջ աշխարհում: Բոլոր բիզնես օբյեկտները, ներառյալ խանութները, շուկաները, վճարային համակարգերը, բանկերը, հսկայական քանակությամբ տվյալներ են կուտակում: Դրանցից օգուտ քաղելու, միտումները, օրինաչափությունները բացահայտելու, կանխատեսումներ անելու և բիզնես ու ֆինանսական մոդելներ ստեղծելու համար այս բոլոր տվյալները պետք է հավաքվեն, համակարգվեն, վերլուծվեն և մեկնաբանվեն: Այս ամենը տվյալների վերլուծության հիմնական խնդիրն է: Python-ը ոչ միայն ունի վերսալ է, այլև պարզ գործիք է, որն օգնում է լուծել գրեթե ցանկացած վերլուծական խնդիր:

Տնտեսագիտական վերլուծության մեջ Python-ի օգտագործումը հեշտացնում է մի շարք գործողություններ տվյալների հետ, որոնցից են՝

- տվյալների հավաքագրման ավտոմատացում;
- մեծացնում է տվյալների վերլուծության արագությունը
- բոլորովին նոր մոտեցումները վեր է ածում վերլուծության :

Այսպիսով, Python-ի օգտագործման շնորհիվ հնարավոր դարձավ լուծել որոշակի խնդիրներ՝ օգտագործելով նյարդային ցանցերի ուսուցումը:

### **Գրականության ցանկ**

1. Պոլ Բեկոն 2018թ. «Ինչու՞ են տնտեսագետները սովորում ծրագրավորել»: Ինտերնետ հասցե [www.refinitiv.ru](http://www.refinitiv.ru)
2. Ս.Ա. Կասպերովիչ 2007թ. «Կանխատեսումը և պլանավորումը տնտեսագիտության մեջ» 173էջ, Ինտերնետ հասցե <https://www.rea.ru>
3. Tadviser ընկերություն 2020թ. «Մեծ տվյալներ: Միջազգային շուկա»
4. Python Software Foundation 2021թ. «Ուղեցույց Python-ի համար» Ինտերնետ հասցե <https://wiki.python.org/moin/BeginnersGuide>
5. Pandas development team 2021թ. «Ուղեցույց Pandas-ի համար» Ինտերնետ հասցե [https://pandas.pydata.org/docs/user\\_guide/index.html](https://pandas.pydata.org/docs/user_guide/index.html)
6. RUVDS ընկերություն 2020թ. «Կարճ ներածություն Pandas-ի մասին», 20էջ Ինտերնետ հասցե <https://habr.com/ru/company/ruvds/blog/494720/>
7. Կուզնեցով 2013թ. «Տվյալների վերլուծություն Pandas-ի միջոցով» 5էջ
8. Ինտերնետ հասցե <https://habr.com/ru/post/196980/>

**Gevorg GRIGORYAN, Aharon RUSHANYAN**

**Analyzing data in economics using a statistical analytics package Python**

*Key words: economic analysis, Python, data analysis, Pandas, economic forecast, economic model*

Recently, the Python programming language has been actively used for data analysis in various socially significant areas. It is one of the most widely used modern programming languages for data analysis. This is mainly due to the simplicity of the language, as well as the availability and variety of modern libraries. The purpose of this paper is to demonstrate the simplicity of processing large amounts of data using Python tools. This article provides an example of classification and exploration of unstructured data, and also to develop an economic forecasting model using Python tools. The article proposes to use the capabilities of Python as a tool that provides high efficiency and accuracy in the use of economical and statistical methods of data analysis.

**Ահարոն ՌՈՒՇԱՆՅԱՆ, Գևորգ ԳՐԻԳՈՐՅԱՆ**

**Տվյալների վերլուծությունը տնտեսագիտության մեջ Python**

**վիճակագրական վերլուծական ծրագրային փաթեթի օգտագործմամբ**

*Բանալի բառեր. տնտեսագիտական վերլուծություն, Python, տվյալների վերլուծություն, Pandas, տնտեսագիտական կանխատեսում, տնտեսագիտական մոդել*

Python ծրագրավորման լեզուն վերջերս ակտիվորեն օգտագործվում է տվյալների վերլուծության համար՝ սոցիալապես նշանակալի տարբեր ոլորտներում: Այն ժամանակակից ծրագրավորման ամենատարածված լեզուներից մեկն է, որը լայնորեն օգտագործվում է տվյալների վերլուծության մեջ: Դա հիմնականում պայմանավորված է լեզվի պարզությամբ, ինչպես նաև ժամանակակից գրադարանների առկայությամբ և բազմազանությամբ: Աշխատանքի նպատակն է ցույց տալ Python գործիքների օգտագործմամբ բավականին մեծ քանակությամբ տվյալների մշակման պարզությունը: Ներկայացված է ոչ կառուցվածքային տվյալների ուսումնասիրության և դասակարգման օրինակ:

**Аарон РУШАНЯН, Геворг ГРИГОРЯН**

**Анализ данных в экономике с помощью пакета программ статистической аналитики Python**

*Ключевые слова: экономический анализ, Python, анализ данных, экономический прогноз*

В последнее время язык программирования Python активно используется для анализа данных в различных социально значимых областях. Это один из наиболее широко используемых современных языков программирования при анализе данных. В основном это связано с простотой языка, а также с доступностью и разнообразием современных библиотек. Цель данной работы - продемонстрировать простоту обработки больших объемов данных с помощью инструментов Python. В этой статье приводится пример классификации и исследования неструктурированных данных, а также разработать модель экономического прогнозирования с использованием инструментов Python. В статье предлагается использовать возможности Python как инструмента, обеспечивающего высокую эффективность.