

емес) құруды жүзеге асыратын процедура деп аталады. Бит операциялары әдетте ығысуды білдіреді.

*Қолданылған әдебиеттер*

1. Информатика. Базовый курс: учебное пособие для высших технических учебных заведений / [С. В. Симонович и др.]. – Санкт-Петербург: Питер, 2017. – 639 с.
2. Иопа, Н. И. Информатика: (для технических специальностей): учебное пособие / Н. И. Иопа. – Москва: КноРус, 2016.
3. Васильков, А.В. Информатика: Учебное пособие / А.В. Васильков, А.А. Васильков, И.А. Васильков. - М.: Форум, 2017.

*Жанат Авдулазим*

*3 курс студенті, академик Е.А. Бөкетов атындағы Қарағанды университеті*

*Турсынғалиева Г.Н.*

*аға оқытушы, академик Е.А. Бөкетов атындағы Қарағанды университеті*

## **BEAUTIFULSOUP ПАКЕТІНІҢ КӨМЕГІМЕН ДЕРЕКТЕРДІ ТАЛДАУҒА АРНАЛҒАН АҚПАРАТТЫ ЖИНАУ**

Кез келген деректерді талдаудың мақсаты – зерттелетін жағдайды толық түсіну (тенденцияларды, оның ішінде жоспардан теріс ауытқуларды анықтау, болжау және ұсыныстар алу). Осы мақсатқа жету үшін деректерді талдаудың келесі міндеттері қойылады:

- ақпарат жинау,
- ақпаратты құрылымдау,
- заңдылықтарды анықтау, талдау,
- болжау және ұсыныстар алу.

Талдау жасап, әдемі графиктерді құрастырмас бұрын, ақпаратты жинау алғашқы сатылардың бірі екендігі анық .

Кәзіргі таңда веб-сайттардан ақпаратты жинаудың түрлі тәсілдері өтек көп. Солардың бірі – сайтты парсингтеу. Бұл веб-сайттардан ақпаратты алу әдісі және бірінші кезекте құрылымдалмаған деректерді - HTML пішіміндегі - вебтегі

құрылымдық деректерге: дерекқорларға немесе электрондық кестелерге түрлендіруге бағытталған. Веб-сайтты парсингтеу HTTP арқылы немесе веб-шолғыш арқылы Интернетке тікелей кіруді қамтиды. Веб-деректерді шығара алатын әртүрлі тілдерде көптеген кітапханалар мен фреймворктар бар болса да, Python көптеген веб-скрапинг мүмкіндіктеріне байланысты көп қолданылады.

Қажетті ақпаратты жинау үшін Python-ды пайдаланып, келесі әрекеттерді орындаймыз:

- Деректерді шығарғымыз келетін беттің URL мекенжайын алу;

- Беттің HTML мазмұнын көшіру немесе жүктеу;
- HTML мазмұнын талдау және қажетті деректерді алу.

Бұл реттілік қалаған беттің URL-мекенжайына өтуге, HTML мазмұнын алуға және қажетті деректерді талдауға көмектеседі. Бірақ кейде деректерді алу үшін алдымен сайтқа кіріп, содан кейін белгілі бір мекенжайға өту керек. Бұл жағдайда сайтқа кіру үшін тағы бір қадам қосылады.

HTML мазмұнын талдау және қажетті деректерді алу үшін BeautifulSoup кітапханасы пайдаланылады. Бұл - HTML және XML құжаттарын парсингтеуге арналған Python пакеті. Ал, дұрыс парсингтеу үшін сайттардың құрылымын түсіну керек. Олардың барлығы дерлік HTML тілі арқылы жасалған.

Қазіргі таңда COVID-19 коронавирусының таралуы жайлы болжамдар өте көп жасалуда. Бұл болжамдарды жасамас бұрын, қажетті ақпаратты интернет көзінен жинау қажет. Алдымен, болжам мен талдау үшін инфекциялар, өлім және сауығулар туралы тарихи деректер керек. COVID-19 вирусы туралы деректер <https://www.worldometers.info/coronavirus/countries-where-coronavirus-has-spread/> бетте еркін қол жетімді.

Ақпаратты сайттан алмас бұрын Python программалау ортасына «requests», «bs4» и «texttable кітапханалары орнатылуы тиіс.

Қажетті кітапханаларды орнатқан соң, келесі қадам - BeautifulSoup пакеті арқылы қажетті кодты жазып, әртүрлі елдердегі жаңа коронавирустың (COVID-19) расталған, өлім-жітім, сауығып кеткен және белсенді жағдайлары туралы соңғы деректерді алу (1-сурет).

```

1 import requests
2 from bs4 import BeautifulSoup
3
4 url = 'https://www.worldometers.info/coronavirus/countries-where-coronavirus-has-spread/'
5
6 page = requests.get(url)
7 soup = BeautifulSoup(page.text, 'html.parser')
8
9 data = []
10
11 data_iterator = iter(soup.find_all('td'))
12
13 while True:
14     try:
15         country = next(data_iterator).text
16         confirmed = next(data_iterator).text
17         deaths = next(data_iterator).text
18         continent = next(data_iterator).text
19
20         data.append((country, int(confirmed.replace(',','')), int(deaths.replace(',','')), continent))
21

```

1-сурет. BeautifulSoup пакетін импорттау және , берілген URL-мекенжайға қосылу

Деректерді оқи алатын форматта көрсету үшін texttable кітапханасын қолданамыз (2-сурет).

```

data.sort(key = lambda row: row[1], reverse = True)
import texttable as tt
table = tt.Texttable()

table.add_rows([(None, None, None, None)] + data)

table.set_cols_align(('c', 'c', 'c', 'c'))
table.header((' Country ', ' Number of cases ', ' Deaths ', ' Continent '))

print(table.draw())

```

2-сурет. Texttable кітапханасын қолдану

Нәтижесінде әр мемлекеттің короновирус жайлы ақпаратты жүктелді. Осы ақпаратты пайдаланып әрі қарай талдаулар жасап, оны визуалдауға болады (3-сурет).

Slovakia	1546510	18817	Europe
Pakistan	1517512	30298	Asia
Norway	1341163	1753	Europe
Ireland	1332701	6590	Europe
Kazakhstan	1304220	13642	Asia
Morocco	1161925	16033	Africa
Bulgaria	1109581	36028	Europe

3-сурет. Веб-сайттан жүктелген деректер

Алынған деректерді \*.txt, \*.csv келтіруші құжаттар ретінде сақтап алуға болады(4-сурет).

Country	Number of cases	Deaths	Continent
United States	80446580	969602	North America
India	42894345	513258	Asia
Brazil	28580995	647486	South America
France	22534971	137770	Europe
United Kingdom	18773164	161104	Europe
Russia	15928568	348578	Europe
Germany	14311052	122895	Europe
Turkey	13841889	93539	Asia
Italy	12651251	154013	Europe
Spain	10949997	99162	Europe
Argentina	8878486	125872	South America
Iran	7011932	135952	Asia
Netherlands	6224073	21671	Europe

4-сурет. \*.txt келтіруші файл

Келесі код көмегімен еліміздегі статистиканы веб-саттан алуға болады (5-сурет):

```
>>> import requests
>>> from bs4 import BeautifulSoup
>>> url = "https://www.worldometers.info/coronavirus/country/kazakhstan/"
>>> req = requests.get(url)
>>> bsObj = BeautifulSoup(req.text, "html.parser")
>>> data = bsObj.find_all("div", class_ = "maincounter-number")
>>> print("Всего заболевших: ", data[0].text.strip())
Всего заболевших: 1,304,220
>>> print("Всего смертей: ", data[1].text.strip())
Всего смертей: 13,642
>>> print("Всего выздоровевших: ", data[2].text.strip())
Всего выздоровевших: 1,279,441
```

5-сурет. Қазақстандағы коронавирус бойынша статистика

BeautifulSoup requests қосымшасының көмегімен жұмыс істейді. Ол қажетті сайттарға html сұрауларын ұйымдастырады, ал BeautifulSoup арқылы алынған ақпарат өңделінеді.

*Жанғали Б.Е.  
4 курс студенті, академик Е.А. Бөкетов атындағы Қарағанды  
университеті  
Хасенова А.А.  
аға оқытушы, академик Е.А. Бөкетов атындағы Қарағанды  
университеті*

## **КИНОФИЛЬМДЕРДІ КӨРУГЕ АРНАЛҒАН МОБИЛЬДІ ҚОСЫМША**

Технологияның дамуы біздің өмірімізді мобильді құрылғылармен тығыз байланыстырды, олардың рөлін бүгінде бағаламау қиын.

Мобильді қосымша - белгілі бір платформаға арналған смартфондарда, планшеттерде және басқа да мобильді құрылғыларда жұмыс істеуге арналған бағдарламалық жасақтама.

Мобильді қосымшалар нарығы бүгінде өте дамыған және үнемі өсіп келеді. Statista болжамдары бойынша, 2020 жылы