

Федеральное государственное бюджетное образовательное учреждение высшего образования
«Сибирский государственный университет телекоммуникаций и информатики»
(СибГУТИ)

09.03.01 "Информатика и вычислительная техника"
профиль "Программное обеспечение средств
вычислительной техники и автоматизированных систем"

ОТЧЕТ

по учебной практике
на кафедре Прикладной Математики и Кибернетики

Выполнил:

студент гр.ИП-811

_____/Разумов Д. Б./

«04» июля 2020г.

Руководитель практики

доцент каф. ПМиК

_____/Приставка П.А./

«04» июля 2020г.

Оценка _____

Новосибирск 2020г.

Содержание

Условие задачи.....	3
Описание программы.....	4
Листинг программы.....	5
Скриншоты.....	7
Источники.....	8

Условие задачи

1. Язык разработки:

Python версии 3.8

2. Операционная система:

Windows 10

3. Набор свойств криптовалют:

- Name – наименование
- Market Cap – рыночная капитализация
- Price – стоимость 1 ед. в долларах США (USD)

4. Ввод данных:

Непосредственно с главной страницы сайта coinmarketcap.com в момент запуска программы. Загрузка и парсинг веб-страницы производится с помощью библиотек Requests и BeautifulSoup или их аналогов

5. Хранение. Типы и структуры для хранения данных

6. Обработка. Реализовать функцию поиска информации о свойствах криптовалюты по ее названию.

Описание программы

1. Функция `get_html(url, params=None)` получает DOM-дерево по адресу `url` с помощью `get-запроса` из пакета `requests`. Дополнительных параметров по умолчанию нет. Функция возвращает полученный контент.
2. Функция `get_html_content(html)` создает класс пакета `BeautifulSoup`, с помощью которого создается коллекция объектов `items`. Эта коллекция состоит из той части `html`-кода сайта, которая содержит информацию о криптовалютах. Далее создается словарь для криптовалют с именем `currencies`. В цикле для объектов `items` ищется информация по названию криптовалюты, его капитализации и стоимости. Эта информация сохраняется в `currencies`. Информация ищется с помощью методов `find()`, `find_all()`. Первый параметр этих методов - это вид тега (`<a>`, `<td>`, `` и т. д.), второй параметр - название нужного класса, которое можно узнать, изучая код сайта. Функция возвращает `currencies`.
3. Функция `save_file(items, path)` сохраняет полученные данные (`items`) о криптовалютах в виде файла `.csv` с именем `path`. Для этого использован пакет `csv`. С помощью полученного файла можно наглядно понять, какие криптовалюты были получены.
4. Функция `search_info(items)` ищет в `items` заданное название криптовалюты и выводит его свойства.
5. Функция `parse()` вызывает все вышеуказанные функции.
6. `URL`, `HEADERS`, `FILE`, `COUNT_MAX` - это условные константы, которые используются в функциях. `URL` - это адрес сайта `coinmarketcap.com`. `HEADERS` - это заголовки, которые нужны для того, чтобы сервер сайта не посчитал программу за бота и не ограничил доступ. `FILE` - это название файла `currencies.csv`, в котором представлены все `COUNT_MAX` (то есть 25) валют.

Листинг программы

```
import requests #использование пакетов
from bs4 import BeautifulSoup
import csv

URL = 'https://coinmarketcap.com/'
HEADERS = {'user-agent': #заголовки, чтобы сервер не посчитал за бота
            'Mozilla/5.0 (Windows NT 10.0; Win64; x64; rv:75.0) Gecko/20100101 Firefox/75.0',
            'accept': '*//*'}
FILE = 'currencies.csv'
COUNT_MAX = 25

def get_html(url, params=None):
    r = requests.get(url, headers=HEADERS, params=params) #get-запрос
    return r

def get_content(html):
    soup = BeautifulSoup(html, 'html.parser') #создание класса, второй параметр - тип
    документа
    items = soup.find_all('tr', class_='cmc-table-row') #коллекция объектов с указанными
    аргументами
    currencies = [] #словарь для криптовалют
    count = 0
    for item in items:
        currencies.append({
            'name': item.find('div', class_='cmc-table__column-name sc-1kxikfi-0
eTVhdN').get_text(),
            'market_cap': item.find('td', class_='cmc-table__cell ' +
                                   'cmc-table__cell--sortable cmc-table__cell--right ' +
                                   'cmc-table__cell--sort-by__market-cap').get_text(),
            'price': item.find('td', class_='cmc-table__cell cmc-table__cell--sortable ' +
                               'cmc-table__cell--right cmc-table__cell--sort-by__price').get_text(),
        })
        count += 1
        if count == COUNT_MAX:
            break
    #print(currencies)
    #print(len(currencies))
    return currencies

def save_file(items, path):
    with open(path, 'w', newline='') as file: #открытие с автоматическим закрытием
        writer = csv.writer(file, delimiter=';') #delimiter - разделитель, чтобы открывалось в
        экселе
        writer.writerow(['name', 'market cap', 'price']) #записали строку
        for item in items:
            writer.writerow([item['name'], item['market_cap'], item['price']])

def search_info(items):
    count = 0
    while 1:
        name = input("\nВведите название криптовалюты, для выхода введите q:\n")
```

```

if name == 'q':
    return
for item in items:
    if item.get('name') == name:
        print("рыночная капитализация:", item.get('market_cap'))
        print("стоимость 1 ед. в долларах США:", item.get('price'))
        break
    count += 1
    if count == COUNT_MAX:
        print('Данное название не найдено')
        count = 0

def parse():
    html = get_html(URL)
    if html.status_code == 200: #если удачная передача контента
        currencies = get_content(html.text)
        save_file(currencies, FILE)
        print('Парсинг завершен')
        search_info(currencies)
    else:
        print('Error')

parse()

```

Скриншоты

```
E:\programming\C4_python>python parser.py
Парсинг завершен

Введите название криптовалюты, для выхода введите q:
Bitcoin
рыночная капитализация: $162,292,456,854
стоимость 1 ед. в долларах США: $8,838.84

Введите название криптовалюты, для выхода введите q:
bitcoin
Данное название не найдено

Введите название криптовалюты, для выхода введите q:
TRON
рыночная капитализация: $1,081,185,893
стоимость 1 ед. в долларах США: $0.016214

Введите название криптовалюты, для выхода введите q:
OLKOKJM
Данное название не найдено

Введите название криптовалюты, для выхода введите q:
q

E:\programming\C4_python>_
```

	A	B	C	D
1	name	market cap	price	
2	Bitcoin	\$162,292,456,854	\$8,838.84	
3	Ethereum	\$22,882,562,180	\$206.54	
4	XRP	\$9,636,731,049	\$0.218456	
5	Tether	\$6,400,476,663	\$1.01	
6	Bitcoin Cash	\$4,501,593,730	\$244.68	
7	Bitcoin SV	\$3,764,746,472	\$204.64	
8	Litecoin	\$3,038,112,716	\$46.99	
9	Binance Coin	\$2,641,961,673	\$16.99	
10	EOS	\$2,554,084,731	\$2.77	
11	Tezos	\$1,990,029,869	\$2.81	
12	Stellar	\$1,475,284,388	\$0.072552	
13	Chainlink	\$1,316,779,489	\$3.76	
14	Cardano	\$1,286,960,276	\$0.049638	
15	TRON	\$1,081,185,893	\$0.016214	
16	UNUS SED LEO	\$1,068,453,438	\$1.07	
17	Monero	\$1,061,841,586	\$60.51	
18	Crypto.com Coin	\$1,003,038,787	\$0.060412	
19	Huobi Token	\$924,116,592	\$4.15	
20	Ethereum Classic	\$839,151,522	\$7.21	
21	Dash	\$753,204,321	\$79.48	
22	USD Coin	\$709,215,675	\$1.00	
23	Neo	\$639,159,254	\$9.06	
24	HedgeTrade	\$622,132,155	\$2.16	
25	Cosmos	\$528,237,960	\$2.77	
26	IOTA	\$513,274,836	\$0.184662	
27				

Источники

1. Официальная документация Python [Электронный ресурс]. URL: <https://docs.python.org/3/tutorial/index.html> (дата обращения: 28.03.2020).
2. Самоучитель Python 3 [Электронный ресурс]. URL: <https://pythonworld.ru/samouchitel-python> (дата обращения: 28.03.2020).
3. Электронная книга из библиотеки СибГУТИ [Электронный ресурс]. URL: <http://www.iprbookshop.ru/52211.html> (дата обращения: 28.03.2020)
3. Работа с веб-данными с помощью Requests и Beautiful Soup в Python 3 [Электронный ресурс]. URL: <https://www.8host.com/blog/rabota-s-veb-dannymi-s-pom..> (дата обращения: 28.03.2020).
4. Web Scraping страниц с помощью Beautiful Soup и Python 3 [Электронный ресурс]. URL: <https://www.8host.com/blog/web-scraping-stranic-s-pom..> (дата обращения: 28.03.2020).