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-011 \_\_\_\_\_\_\_\_\_\_\_\_ \_/Велижанин А.А/

ФИО студента

«28» мая 2022г.

Руководитель практики

доцент каф. ПМиК \_\_\_\_\_\_\_\_\_\_\_\_\_/Приставка П.А./

«28» мая 2022г. Оценка\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Новосибирск 2022 г.

**Условие задачи**

Разработать программу реализующую ввод, хранение и обработку данных о котировках криптовалют на основе данных сайта coinmarketcap.com.

Требования:

1. Язык разработки: Python версии не ниже 3.x
2. Операционная система: определяются студентом
3. Набор свойств криптовалют:

Name – название крипты.

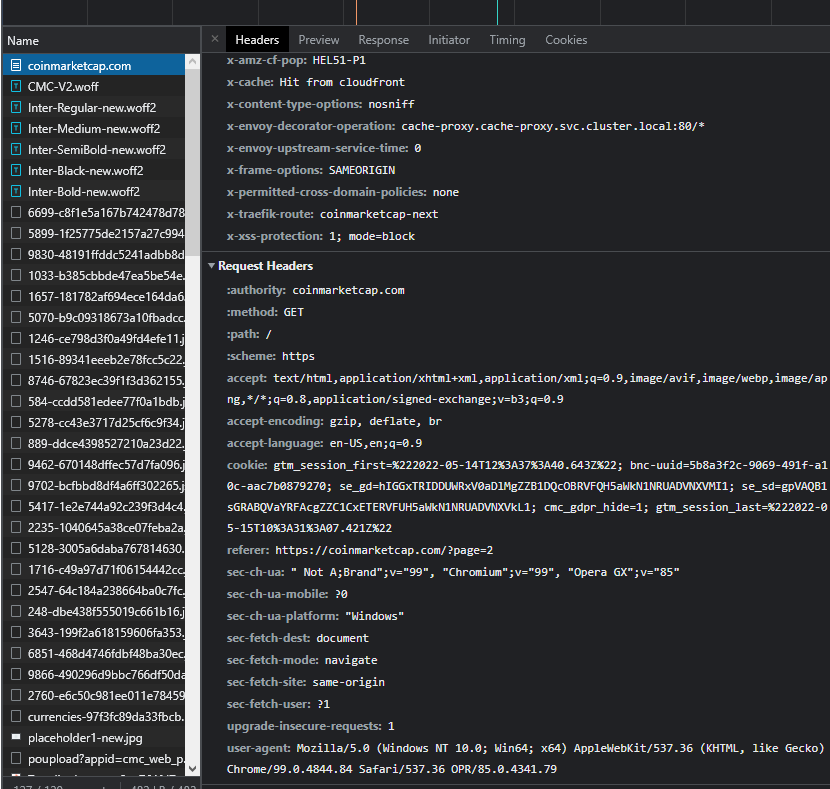
Market\_cap – капитализация.

Price- стоимость в usd.

Описание алгоритмов:

def search\_info – Ну тут я ищу валюту по ее названию.

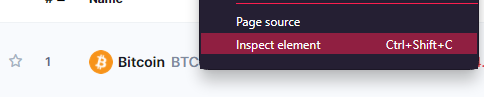
User-Agent взял из вкладки network в доке 





Вообще при необычайно большом количестве запросов юзер улетает в бан.

Поэтому лучше при парсинге других сайтов сделать несколько кастомных и рандомить между ними для минимизации шанса бана.

При поиске данных для файлика надо было шарить по элементам сайта 

У каждой крипты есть данные: название, цена и т.д.

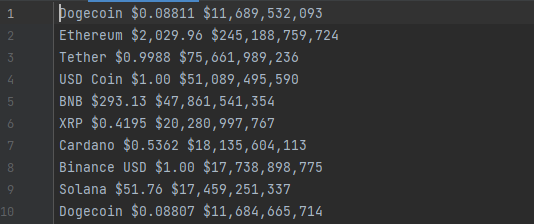
У них есть свои отличительные черты например: div с классом sc-16r8icm-0 sc-1teo54s-1 dNOTPP .

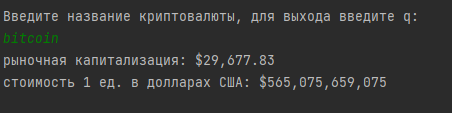
Именно по ним и происходил сбор данных или правильно сказать парсинг ?

Листинг программы:

import requests  
from bs4 import BeautifulSoup as bs  
  
  
def write\_txt(data, f\_name='new\_file.txt'):  
 with open(r"C:\Users\189gl\PycharmProjects\pythonProject\new\_file.txt", "a") as file:  
 file.write(data + '\n')  
 file.close()  
 return f\_name  
  
  
def search\_info():  
 count = 0  
 while 1:  
 name = input('\nВведите название криптовалюты, для выхода введите q:\n')  
 if name == 'q':  
 return  
 for item in open('C:\\Users\\189gl\\PycharmProjects\\pythonProject\\new\_file.txt'):  
 data\_fak = item.split()  
 if data\_fak[0].lower() == name.lower():  
 print("рыночная капитализация:", data\_fak[1])  
 print("стоимость 1 ед. в долларах США:", data\_fak[2])  
 break  
 count += 1  
  
  
url = 'https://coinmarketcap.com/'  
headers = {'User-Agent': 'Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) '  
 'Chrome/99.0.4844.84 Safari/537.36 OPR/85.0.4341.79 '  
 }  
  
r = requests.get(url, headers=headers)  
soup = bs(r.text, 'lxml')  
trs = soup.find('tbody').find\_all('tr', limit=10)  
name\_f = []  
price\_f = []  
market\_cup\_f = []  
for tr in trs:  
 quotes = tr.find\_all('div', class\_='sc-16r8icm-0 escjiH')  
 names = tr.find\_all('p', class\_='sc-1eb5slv-0 iworPT')  
 prices = tr.find\_all('div', class\_='sc-131di3y-0 cLgOOr')  
 market\_cups = tr.find\_all('span', class\_='sc-1ow4cwt-1 ieFnWP')  
 for name in names:  
 name\_f.append(name.text)  
 for price in prices:  
 price\_f.append(price.text)  
 for market\_cup in market\_cups:  
 market\_cup\_f.append(market\_cup.text)  
print(name\_f)  
for i in range(0, 10):  
 print(name\_f[i] + " " + price\_f[i] + " " + market\_cup\_f[i])  
 write\_txt(name\_f[i] + " " + price\_f[i] + " " + market\_cup\_f[i])  
  
if \_\_name\_\_ == "\_\_main\_\_":  
 search\_info()

Результаты:





Использованные источники:

https://pythonru.com/biblioteki/parsing-na-python-s-beautiful-soup