

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное автономное образовательное учреждение высшего
образования
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ ЯДЕРНЫЙ УНИВЕРСИТЕТ
«МИФИ»

ИНСТИТУТ ФИНАНСОВОЙ И ЭКОНОМИЧЕСКОЙ БЕЗОПАСНОСТИ
КАФЕДРА ФИНАНСОВОГО МОНИТОРИНГА

ОТЧЕТ
ПО ЛАБОРАТОРНОЙ РАБОТЕ №4
НА ТЕМУ:
ИСПОЛЬЗОВАНИЕ ЦИКЛОВ ДЛЯ ИЗВЛЕЧЕНИЯ БОЛЬШИХ ДАННЫХ
ПО КУРСУ «РИАС»

Выполнил: Нуритдинходжаева А. (С20-702).

Преподаватель: Прохоров И.В.

Оценка _____ Дата _____

Москва, 2024

Цель: освоить технологию извлечения данных объемом более 1000 объектов из социальной сети Вконтакте посредством организации циклов на языке Python.

Задание

Подсчитать программой на языке Python число людей в сообществах.

Для выполнения данной лабораторной работы будем использовать access.token, полученный в лабораторной работе №3: vk1.a.x3wkE8A5yrPoMDRoXtpA7gcCvHh-ybVaPk_YgNQx9VaJF5sA1cWtPqDg3AtTaAwtuDRas1kvizHskNjXtJLpjFXAALdB3dyL7ezkJHm2YmsfIoL0Q6J08Mw-__McGCT5XLSIB5pXnhICkgivUf_XT-dGzbPwRSdpADvVmbiN2cox6p0TpGsXh3QMbcgY1HWPhncJC3Yv1g1hYU3-fPL1CA

По заданию лабораторной работы была написана программа:

```
import time
import requests
import matplotlib.pyplot as plt

access_token = 'vk1.a.x3wkE8A5yrPoMDRoXtpA7gcCvHh-
ybVaPk_YgNQx9VaJF5sA1cWtPqDg3AtTaAwtuDRas1kvizHskNjXtJLpjFXAALdB3dyL7ezkJH
m2YmsfIoL0Q6J08Mw-__McGCT5XLSIB5pXnhICkgivUf_XT-
dGzbPwRSdpADvVmbiN2cox6p0TpGsXh3QMbcgY1HWPhncJC3Yv1g1hYU3-fPL1CA'
groups = ["rtvi", "incrussiamedia", "gso", "molnyja", "fcveles",
"vhl_official", "ice_show", "djsnakefr", "a_onefilms", "kremlinpalace"]

url =
"https://api.vk.com/method/groups.getMembers?access_token={}&group_id={}&v={}"
"

users = {}
version = "5.131"

for group in groups:
    response = requests.get(url.format(access_token, group, version)).json()
    if "error" in response:
        print(f"For {group}: {response['error']['error_msg']}")
        continue
    time.sleep(1)
    users[group] = response['response']['count']
    print(f"For {group}: {users[group]}")

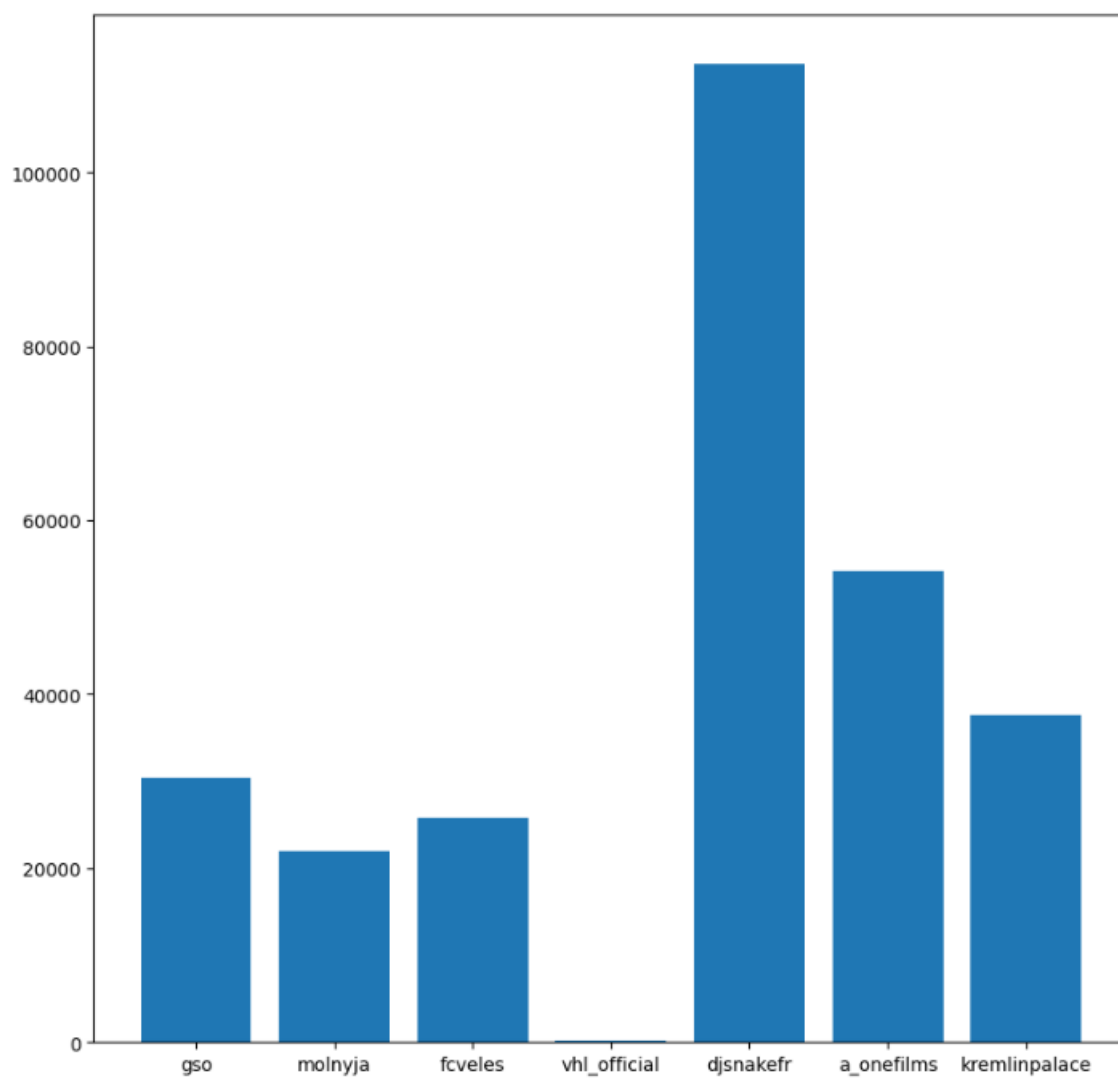
x = users.keys()
plt.figure(figsize=(10, 10))
plt.bar(x, height=[y for y in users.values()])
plt.show()
```

По результатам работы цикла в программе были получены следующие данные о количестве пользователей в группах:

```
For rtvi: Access denied: group hide members
For incrussiamedia: Access denied: group hide members
For gso: 30416
For molnyja: 21902
For fcveles: 25793
For vhl_official: 192
For ice_show: Access denied: group hide members
For djsnakefr: 112508
For a_onefilms: 54127
```

For kremlinpalace: 37630

По полученным данным построена гистограмма распределения количества пользователей по группам:



Заключение

По итогам проведенной лабораторной работы были успешно освоены методы работы с большими данными, использованы сервисы API ВКонтакте для извлечения более 1000 объектов.

При работе с API ВКонтакте были реализованы запросы для получения данных о количестве пользователей в сообществах ВКонтакте. По полученным данным о количестве участников была построена гистограмма распределения количества пользователей по группам.