

Георгий Борисенко
ДОМАШНЕЕ ЗАДАНИЕ
До 12:00 29.08.2026

1 Описание задания

Цель задания - собрать данные из социальной сети ВКонтакте с помощью официального VK API. Необходимо загрузить посты из официальных сообществ крупных новостных изданий.

2 Общие требования

1. Использовать только официальный VK API.
2. Не хранить `access_token` в коде или итоговом архиве.
3. Код должен полностью воспроизводить сбор и сохранение данных.
4. В качестве решения необходимо сдать файл `.ipynb` с рабочим кодом и файл `.parquet` с собранным датасетом.

3 Сообщества

Соберите посты из следующих официальных сообществ:

1. Известия
2. Коммерсант
3. РБК
4. РИА Новости
5. Ведомости

4 Часть 1. Доступ к VK API и организация кода

Подготовьте код для работы с VK API.

1. Получите `access_token` для работы с VK API.
2. Вынесите токен в переменную окружения или отдельный конфигурационный файл.
3. Подготовьте файл `config_example`, который показывает, где должны храниться токен и настройки, но не содержит действующего токена.
4. Реализуйте функцию-обертку для отправки запросов к API.

5 Часть 2. Сбор постов

С помощью метода `wall.get` получите последние 100 постов из каждого указанного сообщества.

Основные параметры метода:

- `domain` - короткое имя сообщества;
- `owner_id` - идентификатор владельца, для сообществ обычно отрицательный;
- `count` - число постов за один запрос, максимум 100;
- `offset` - смещение для пагинации;
- `filter` - тип постов, например `all`.

Для каждого поста сохраните как минимум следующие поля:

- название сообщества;
- идентификатор поста;
- дату публикации;
- текст поста;
- число лайков;
- число комментариев;
- число репостов;
- число просмотров, если оно доступно;
- ссылку на пост, если она доступна.

Сохраните итоговый датасет в файл `posts.parquet`.