

Финальный проект «Собеседование в Зелененький Банк»

4. (*) Парсинг данных по каждому дому с сайта "Реформа ЖКХ". Нужны дома только по Москве и Московской области. Необходимо выгрузить всю информацию, содержащуюся в паспорте дома. На основе этих данных составить таблицу: для каждого адреса сопоставить данные из паспорта. По ходу выполнения задания, можете ответить на вопрос: "Какие способы обхода капчи Вы знаете?".

Результатом этого задания ожидается: 1) код, 2) выгруженная база данных.

Рис. 1: Так выглядело оригинальное условие задания

Оригинальный текст задания приведен на рисунке 1. И также предлагалось выполнять это задание при помощи языка программирования R.

Вам это надо делать на Python и уметь автоматически выгружать информацию только об одном указанном пользователем доме. Выгружать надо только паспорт дома. Предполагается, что вы оформите свое решение в виде модуля.

За это задание можно набрать 40 баллов: 10 баллов за работающее решение, 10 баллов за эффективное использование ООП и других особенностей языка, 10 баллов за документацию к вашему модулю и примеры работы с ним, 10 баллов за продуманную архитектуру модуля (удобное и гибкое взаимодействие классов, позволяющее безболезненно добавлять новый функционал, быстро удовлетворяя новые требования заказчика модуля).

Библиотеки, которые вам могут понадобиться: `urllib`, `lxml`, `pandas`, `beautifulsoup`, `csv`.