

Тестовое задание: Парсинг и работа с данными

ITMO: NSS lab

Сентябрь 2022

1 Парсинг данных и работа с данными

Выберете любой файл Excel для парсинга.

По выбранному файлу необходимо выгрузить календарь работ и календарь ресурсов. То есть для каждой работы и ресурса надо выгрузить, сколько было сделано (использовано) в каждый день месяца. Также для каждой работы и ресурса есть плановое значение расхода, есть фактическое, нужно выгружать оба.

Требования к парсингу:

1. Названия работ и названия ресурсов должны иметь постфикс типа ast для работ и _res для ресурсов;
2. Если файл содержит информацию о субподрядчике она должна быть извлечена.
3. Результат должен быть представлен в виде .csv файла.

Основные проблемы связаны с тем, что в каких-то файлах работы и ресурсы собраны в одном листе, в каких-то файлах это разные листы, а иногда это даже разные файлы.

2 Варианты Файлов для выбора

1. Файл 1 – выгрузить только Ноябрь, работы и ресурсы на одном листе. Доп. проблема – ячейки с датами имеют формат не date, а просто число;
2. Файл 2 – выгрузить все месяцы на вкладке МСГ, ресурсы брать из той жекладки МСГ;
3. Файл 3 – работы выгрузить из листа МСГ, ресурсы из листа Ресурсы. Доп. проблема – формат файла .xlsm;
4. Файл 4 – работы выгрузить из листа СМГ, ресурсы – из листа Люди-техника. Доп. проблема – формат файла .xlsb;
5. Файл 5 – работы выгрузить из листа МСГ, ресурсы – из листа Ресурсы;

6. Файл 6 – работы выгрузить из листа СМГ, ресурсы – из листа Люди Техника;
7. Файл 7 – работы выгрузить из листа МСГ, ресурсы – из листа Ресурсы;
8. Файл 8 – выгрузить работы и ресурсы из листа Декабрь Лот 1;
9. Файл 9 – выгрузить работы и ресурсы из листа Декабрь;
10. Файлы 10 и 11 идут вместе. Работы выгружаются из файла 10 за май (вкладка СМГ). Ресурсы также выгружаются за май из файла 11;
11. Выгрузить работы из СМГ вкладки, ресурсы из вкладки Люди Техника.