



Задание №3*

Загрузка данных из внешнего ресурса.

По результатам данного задания вы углубитесь в тонкости работы с кодировками и CSV форматом.

Инструменты, которые пригодятся для выполнения: Apache NiFi, база данных PostgreSQL

Критерии, по которым будет оцениваться задание:

Данные загружены в БД в полном объёме. Данное задание **не** будет оцениваться ревьювером. Тем не менее, эти знания пригодятся для выполнения задания по модулю.

Описание задания

Скачать из сети файл <https://minpromtorg.gov.ru/opendata/7705596339-refusaloftheconclusion/data-20170202-structure-20161019.csv>

Все данные как есть, записать в БД (Postgresql).

Подвох задания в том, что заголовки в файле на русском языке и кодировка файла cp1251 и зачем-то разделитель в конце строк.

Заголовки надо будет использовать на латинице. Рекомендую сделать AVRO схему. Как вариант, можно поменять заголовки процессором ReplaceTextWithMapping.

Создайте таблицу refusaloftheconclusion для этих данных.



Интеграция NiFi с Redis

По результатам данного задания, вы сможете взаимодействовать с базой данных типа ключ-значение Redis (Remote Dictionary Service).

Инструменты, которые пригодятся для выполнения: Apache NiFi, база данных Redis

Описание задания

По [ссылке](#) ознакомьтесь со статьёй до раздела "Redis Sentinel".

Требуется повторить такой же поток, как в приведённой статье.

Но для RedisConnectionPoolService в поле Database Index установить значение 9.

Проверьте себя самостоятельно

1. Необходимо из командной строки redis выполнить команды

```
SELECT 9
```

```
INFO
```

Дополнительная информация:

В данной задаче можно использовать Docker, например, вот с таким yml



```
version: '3'

networks:

  smart:

    driver: bridge

    name: nifi

services:

  redis:

    image: 'bitnami/redis:latest'

    ports:

      - "6379:6379"

    environment:

      - ALLOW_EMPTY_PASSWORD=yes

    networks:

      - smart
```

Обратите внимание: Если вам не удастся скачать файл и будете получать ошибку

```
unable to find valid certification path to requested target
```

то скачайте, пожалуйста, файл любым способом (wget или curl). Заберите процессором GetFile.

Как вам урок?



Изучил, далее >>

Слёрм ©

[+7 \(495\) 248-05-80](tel:+7(495)248-05-80)

[Лицензия №ДЛ-1368 от 22.08.2019](#)

[Политика конфиденциальности](#)

[Публичная оферта](#)

