**Инструкция по развертыванию проекта**

Для того, чтобы развернуть проект понадобится:

1. Docker

2. VSCode

3. DBeaver (см. файл "Инструкция по установке DBeaver")

4. Windows 10 и старше

**Шаг 1**. Создать папку для проекта в удобной директории.

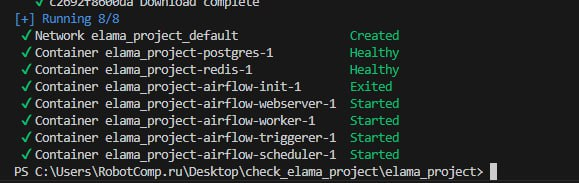
**Шаг 2**. Открыть созданную папку с помощью VSCode. Запустить терминал. В терминале ввести команду docker pull eg0r24/kaggle-repository:kaggle-image-v2025-06-01-r2 , чтобы скачать докер-образ проекта.

**Шаг 3**. В терминале VSCode запустить команду git clone https://github.com/Egor152/elama\_project.git , чтобы скачать код проекта.

**Шаг 4**. Используйте команду cd elama\_project , чтобы перейти в папку проекта.

**Шаг 5**. Введите команду docker compose up -d --build для сборки образа и запуска контейнера, согласно файлу docker-compose.yaml.

**Шаг 6**. Подождите запуска контейнера. Должны запуститься все сервисы:



**Шаг 7**. В браузере в адресной строке введите http://localhost:8080/ - запуститься airflow-webserver. Логин - airflow, пароль - airflow.

**Шаг 8**. Зайдите в Admin - > Connections.

Изображение выглядит как текст, снимок экрана, программное обеспечение

Автоматически созданное описание

Создайте подключение в Airflow с такими параметрами и нажмите "Save". Параметры подключения:

|  |  |
| --- | --- |
| Connection\_id | pg\_connect |
| Connection Type | Postgres |
| Host | postgres |
| Database | airflow |
| Login | airflow |
| Password | airflow |
| Port | 5432 |

Изображение выглядит как текст, снимок экрана, программное обеспечение, число

Автоматически созданное описание

**Шаг 9**. Запустите Dbeaver. Создайте подключение к Postgres.

Для этого нужно нажать на кнопку "Новое соединение" в левом верхнем углу.

Изображение выглядит как текст, программное обеспечение, снимок экрана

Автоматически созданное описание

В меню подключений выбрать "PostgreSQL", щелкнув по нему дважды, либо нажав "Далее".

Изображение выглядит как текст, снимок экрана, программное обеспечение, Операционная система

Автоматически созданное описание

У Вас на экране откроется окно создания подключения. Введите данные со скриншота или таблицы.

Изображение выглядит как текст, снимок экрана, программное обеспечение, число

Автоматически созданное описание

|  |  |
| --- | --- |
| Хост | localhost |
| База данных | airflow |
| Логин | airflow |
| Пароль | airflow |
| Порт | 15432 |

**Шаг 10**. Перейдите в DAGs в UI Airflow. В списке DAG-ов нажмите на "kaggle\_dag". Нажмите кнопку "Trigger DAG" в правом верхнем углу.

Изображение выглядит как текст, снимок экрана, программное обеспечение, веб-страница

Автоматически созданное описание

**Шаг 11**. Дождитесь пока все задачи будут выполнены (станут зелеными).

Изображение выглядит как текст, снимок экрана, Шрифт, число

Автоматически созданное описание

**Шаг 12**. Проверка файлов. В папке проекта по пути elama\_project/data появятся файлы с данными с сайта Kaggle.

Изображение выглядит как текст, снимок экрана, программное обеспечение, Мультимедийное программное обеспечение

Автоматически созданное описание

**Шаг 13**. Зайдите в DBeaver для проверки БД.

В списке слева должно быть подключение к БД "airflow", в ней должна быть схема "stg". В схеме "stg" будет таблица "adv".

Изображение выглядит как текст, снимок экрана, программное обеспечение, веб-страница

Автоматически созданное описание

Вы можете открыть таблицу двойным нажатием на ее название. И у вас откроется таблица "adv":

Изображение выглядит как текст, снимок экрана, число, программное обеспечение

Автоматически созданное описание