

Департамент образования и науки города Москвы
Государственное автономное образовательное учреждение высшего
образования города Москвы
«Московский городской педагогический университет»
Институт цифрового образования
Департамент информатики, управления и технологий

ДИСЦИПЛИНА:

Проектный практикум по разработке ETL-решений

Лабораторная работа №6-1

Тема:

Бизнес-кейс «StockSense»

Выполнил(а): Морозова Валерия АДЭУ-211

Преподаватель:

Москва

2025

Развернута Конфигурация ВМ в VirtualBox.

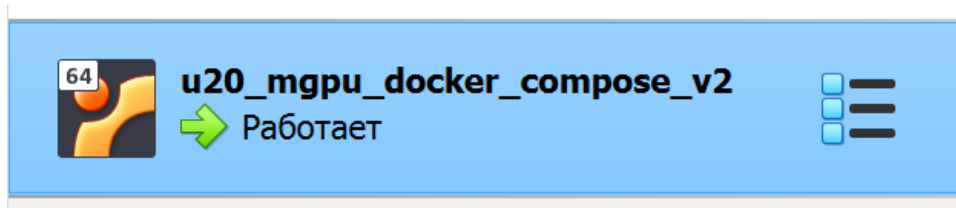


Рисунок 1. Образ развернут

Клонировано на ПК задание Бизнес-кейс «StockSense» в домашний каталог ВМ.

```
git clone https://github.com/BosenkoTM/workshop-on-ETL.git
mgpu@mgpu-VirtualBox:~$ cd Downloads/
mgpu@mgpu-VirtualBox:~/Downloads$ git clone https://github.com/BosenkoTM/workshop-on-ETL.git
Cloning into 'workshop-on-ETL'...
remote: Enumerating objects: 675, done.
remote: Counting objects: 100% (68/68), done.
remote: Compressing objects: 100% (57/57), done.
remote: Total 675 (delta 22), reused 1 (delta 1), pack-reused 607 (from 1)
Receiving objects: 100% (675/675), 5.84 MiB | 3.66 MiB/s, done.
Resolving deltas: 100% (320/320), done.
mgpu@mgpu-VirtualBox:~/Downloads$
```

Рисунок 2. Задание клонировано в домашний каталог

Запуск контейнера с кейсом, изучение основных элементов DAG в Apache Airflow.

Прежде чем запустить контейнер необходимо остановить все контейнеры Docker.

```
● mgpu@mgpu-VirtualBox:~/Downloads/workshop-on-ETL/business_case_stocksense_25$ sudo docker stop $(sudo docker ps -a -q)
47d0965208e9
● mgpu@mgpu-VirtualBox:~/Downloads/workshop-on-ETL/business_case_stocksense_25$ sudo docker rm -f $(sudo docker ps -a -q)
47d0965208e9
● mgpu@mgpu-VirtualBox:~/Downloads/workshop-on-ETL/business_case_stocksense_25$ sudo docker rmi -f $(sudo docker images -q)
Untagged: postgres:12-alpine
Untagged: postgres@sha256:de1426a8865f7f3ca1b030e89529c8c06faa9f288495db0d1b2146805ae391f3
Deleted: sha256:a6a83fb59068527d823cef33c61759f37dfef7b0760cd4e4047a519fe80854
Deleted: sha256:cd33174a169bec29c2063c19d545820fca018469e28e830b5d17e636520ae19b
Deleted: sha256:3c36902f3f690446334f8cbe92b735bea66ae5e22772cc7701dc4095441f73fc
```

Рисунок 3. Остановка и удаление всех контейнеров

```
● mgpu@mgpu-VirtualBox:~/Downloads/workshop-on-ETL/business_case_stocksense_25$ sudo docker build -t custom-airflow:slim-2.8.1-python3.11 .
[+] Building 122.3s (7/7) FINISHED
=> [internal] load build definition from Dockerfile                                docker:default 0.0s
=> => transferring dockerfile: 615B                                              0.0s
=> [internal] load metadata for docker.io/apache/airflow:slim-2.8.1-python3.11 2.0s
=> [internal] load .dockerignore                                                 0.0s
=> => transferring context: 2B                                                  0.0s
=> [1/3] FROM docker.io/apache/airflow:slim-2.8.1-python3.11@sha256:751babd58a83e44ae23c393 59.2s
=> => resolve docker.io/apache/airflow:slim-2.8.1-python3.11@sha256:751babd58a83e44ae23c393 0.0s
=> => sha256:3ee88b8d122ebb0fbb9be864918a05a7621f1b4e1801154b2a0bd64e9476c3 4.47kB / 4.47kB 0.0s
=> => sha256:a205efa96734ac8633bf8d388ed9b6cd527835d31ebee070ba1cedfb880b 25.59kB / 25.59kB 0.0s
=> => sha256:51d1f07906b71fd60ac43c61035514996a8ad8dbfd39d4f570ac5446b064ee 3.51MB / 3.51MB 1.9s
=> => sha256:fe87ad6b112e2dfa9a52f49adf5bb70a80d9af2c737f71a947cb5017a12 12.87MB / 12.87MB 4.8s
=> => sha256:751babd58a83e44ae23c393fe1552196c25f3e2683c97db1a6d98b7d15e7a0 1.61kB / 1.61kB 0.0s
=> => sha256:e1caac4eb9d2ec24aa3618e5992208321a92492aef5fef5eb9e470895f77 29.12MB / 29.12MB 5.5s
```

Рисунок 4. Сборка Docker образа с указанием тэга

```
6bf20047005c56d6fc5b.js HTTP/1.1" 304 0 "http://localhost:8080/dags/listing_4_20/grid
?tab=graph&dag_run_id=manual_2025-04-05T08%3A21%3A44.808427%2B00%3A00" "Mozilla/5.0
(X11; Ubuntu; Linux x86_64; rv:123.0) Gecko/20100101 Firefox/123.0"
webservice-1 | 172.18.0.1 - - [05/Apr/2025:08:22:33 +0000] "GET /object/grid_data?
dag_id=listing_4_20&num_runs=25 HTTP/1.1" 200 13244 "http://localhost:8080/dags/listi
ng_4_20/grid?tab=graph&dag_run_id=manual_2025-04-05T08%3A21%3A44.808427%2B00%3A00" "
Mozilla/5.0 (X11; Ubuntu; Linux x86_64; rv:123.0) Gecko/20100101 Firefox/123.0"
webservice-1 | 172.18.0.1 - - [05/Apr/2025:08:22:33 +0000] "GET /object/next_run_d
atasets/listing_4_20 HTTP/1.1" 200 2 "http://localhost:8080/dags/listing_4_20/grid?ta
b=graph&dag_run_id=manual_2025-04-05T08%3A21%3A44.808427%2B00%3A00" "Mozilla/5.0 (X1
1; Ubuntu; Linux x86_64; rv:123.0) Gecko/20100101 Firefox/123.0"
webservice-1 | 172.18.0.1 - - [05/Apr/2025:08:22:34 +0000] "GET /object/graph_data
?dag_id=listing_4_20 HTTP/1.1" 200 295 "http://localhost:8080/dags/listing_4_20/grid?
tab=graph&dag_run_id=manual_2025-04-05T08%3A21%3A44.808427%2B00%3A00" "Mozilla/5.0 (
X11; Ubuntu; Linux x86_64; rv:123.0) Gecko/20100101 Firefox/123.0"
scheduler-1 | [2025-04-05T08:25:30.794+0000] {scheduler job runner.py:1619} INFO
```

Рисунок 5. Запуск контейнеров

Connection Id * my_postgres

Connection Type * Postgres

Description

Host business_case_stocksense_25-wiki_results-1

Schema airflow

Login airflow

Рисунок 6. Подключение к postgres



Рисунок 7. DAG выполнен успешно

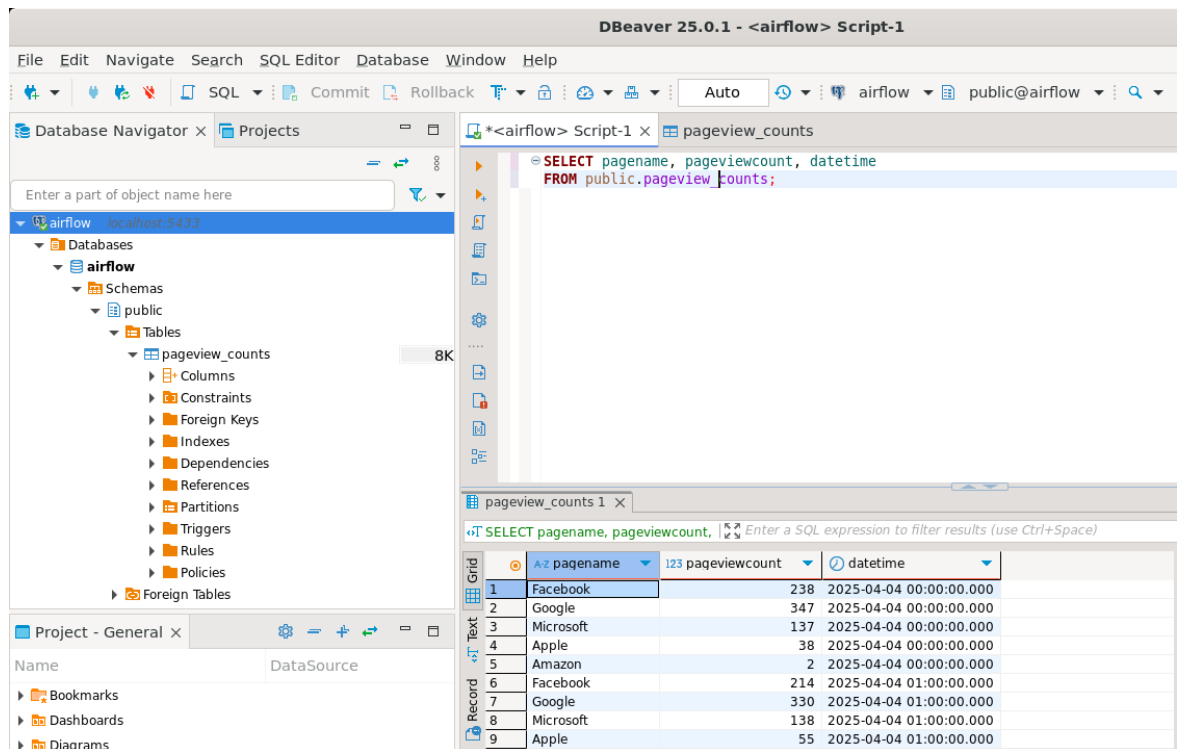


Рисунок 8. Настройка подключения в БД в DBeaver

Как видно, данные успешно выгрузились в таблицу.

6.1.5. Спроектировать верхнеуровневую архитектуру аналитического решения Бизнес-кейса «StockSense» в draw.io. Необходимо использовать:

Source Layer - слой источников данных.

Storage Layer - слой хранения данных.

Business Layer - слой для доступа к данным пользователей.

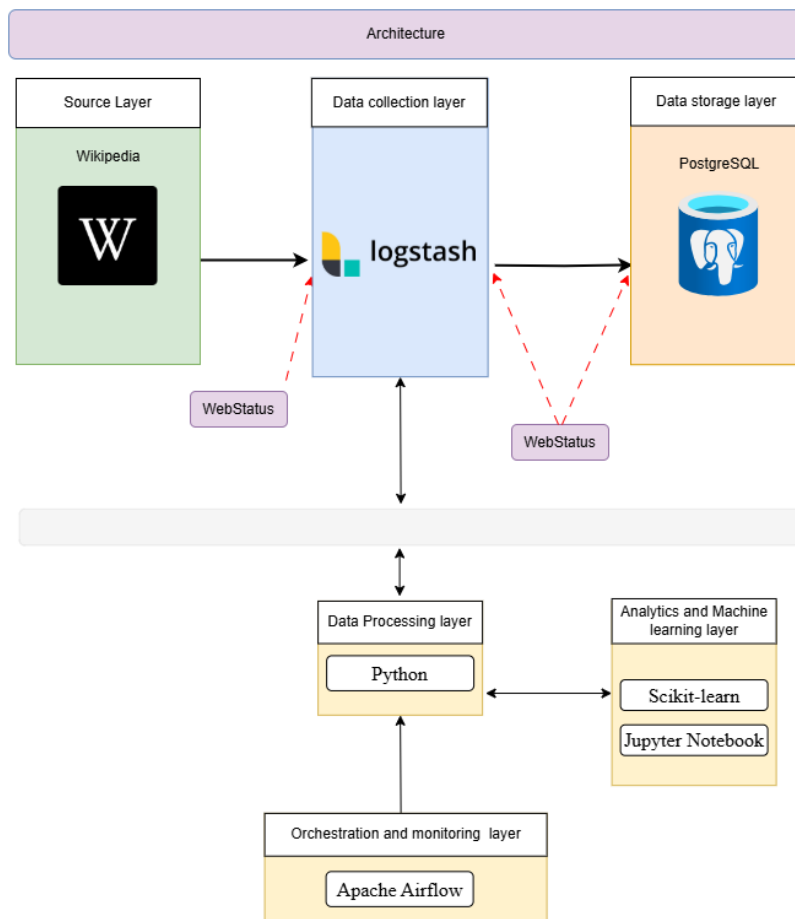


Рисунок 9. Архитектура аналитического решения задания Бизнес-кейса «StockSense»

6.1.6. Спроектировать архитектуру DAG Бизнес-кейса «StockSense» в draw.io. Необходимо использовать:

Source Layer - слой источников данных.

Storage Layer - слой хранения данных.

Business Layer - слой для доступа к данным пользователей.

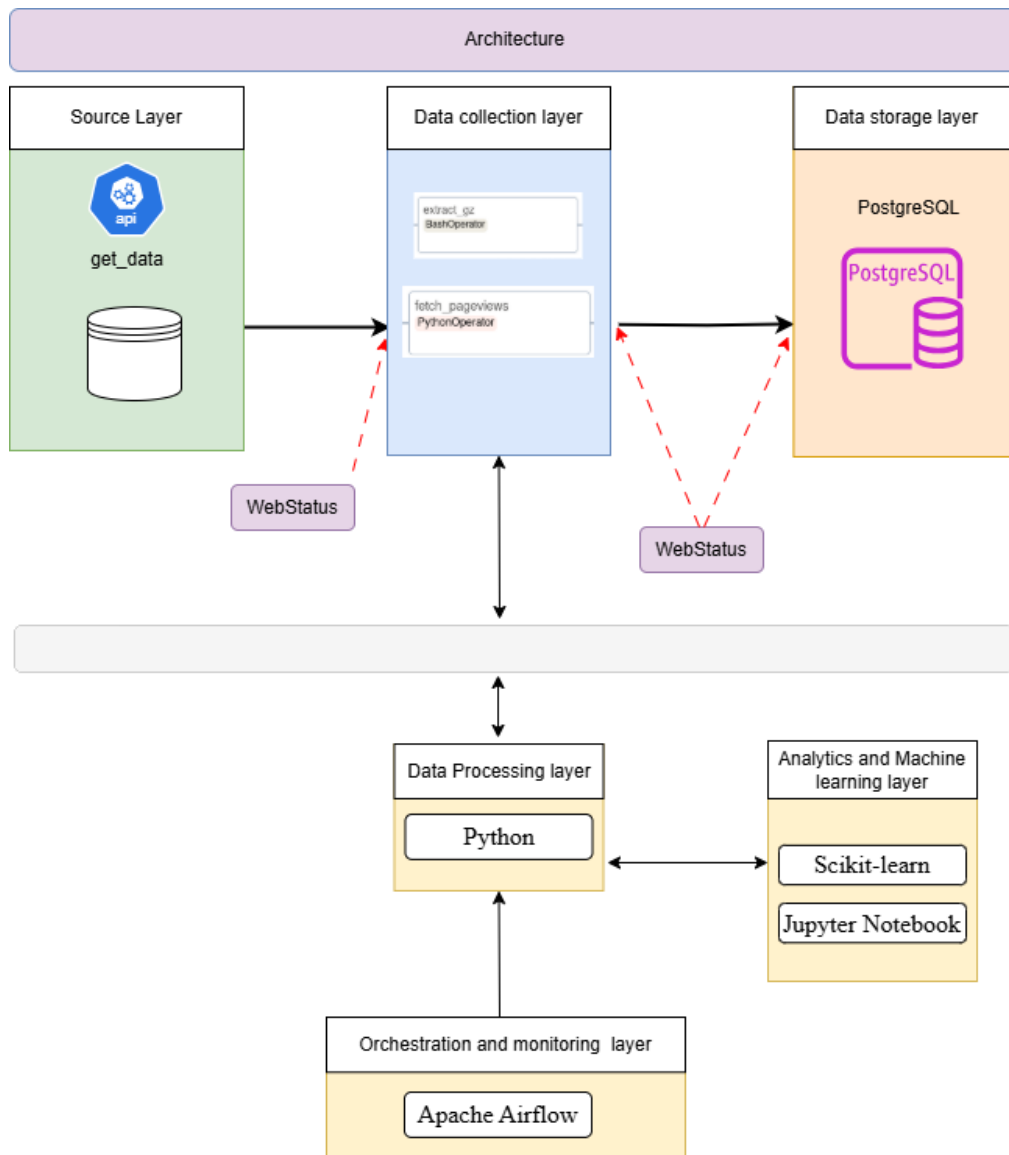


Рисунок 10. Архитектура DAG Бизнес-кейса «StockSense»

6.1.7. Построить диаграмму Ганта работы DAG в Apache Airflow.

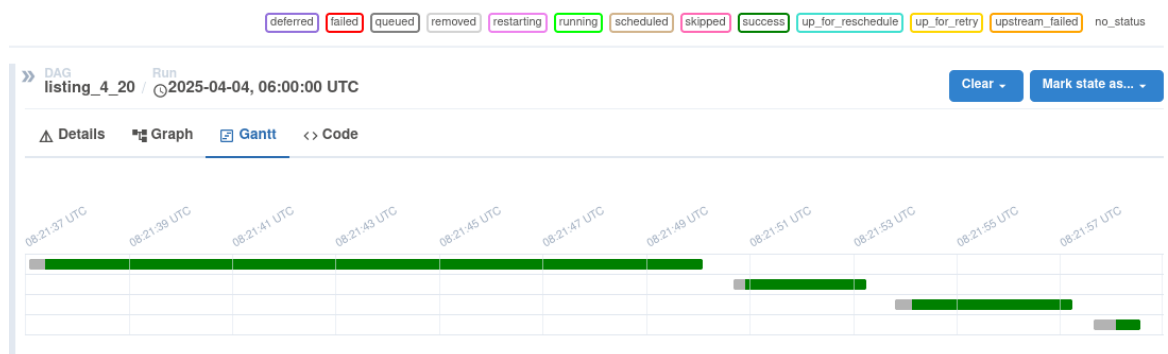


Рисунок 11. Диаграмма Ганта