



## Вариант №1 ETL пайплайны в NoSQL

### Цели практики

Научиться выполнять ETL пайплайн в **Clickhouse**.

### Описание задания:

Задание состоит в том, чтобы научиться выполнять ETL пайплайн в Clickhouse.

- Используя образ Clickhouse, аналогично тому как показано на уроке, необходимо в Docker установить образ сервера БД.
- Для выполнения задания используйте готовый массив данных пользователей для вставки в Clickhouse в CSV формате (файл для скачивания находится на диске [https://disk.yandex.ru/d/YagnK1O\\_qaicOw](https://disk.yandex.ru/d/YagnK1O_qaicOw) ).

### Инструменты, которые пригодятся для выполнения:

- Docker
- docker-compose

### Задание:

#### Создать таблицу в Clickhouse с полями:

- username
- country
- comes\_from
- hobbies
- ip\_address
- birthday

**Заполнить данными из предоставленного файла CSV. Типы данных определить самостоятельно по документации.**

#### Далее составить запросы:

- а) Определить самую популярную страну, откуда пришли пользователи;
- б) Отобразить тех пользователей, которые пришли с домена "mail.ru";
- в) Найти тех пользователей, у которых поле hobbies содержит более 1;

**По итогу выполнения задания у вас должны быть** текстовые файлы с пятью работоспособными запросами и три ответа от БД по трем запросам последнего пункта задания.

### Советы и рекомендации



Запросы рекомендуется писать используя Python. При желании, можно протестировать работу в терминале внутри контейнера.

Для этого зайдите в него командой `sudo docker exec -it <id_container> clickhouse-client` (как было продемонстрировано в уроке).

- Для установки Docker на Windows 10 и 11 воспользуйтесь инструкциями:
- Установить Docker на Windows 10 - <https://treyakov.net/post/ustanovit-docker-na-windows-10-wsl2/>
- Установить Docker на Windows 11 - <https://youtu.be/6k1CyA5zYgg>

Как вам урок?



Далее >

Слёрм ©

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

[Лицензия №ДЛ-1368 от 22.08.2019](#)

[Политика конфиденциальности](#)

[Публичная оферта](#)

