

Задание: Мини-система сбора и анализа данных

Цель: создать простую end-to-end систему, которая генерирует данные, сохраняет их и позволяет анализировать через Redash и Jupyter Notebook.

Компоненты системы

1. Генератор данных (Python-скрипт)

Что делает:

- Создаёт записи с заданной периодичностью (например, раз в секунду)
- Генерирует данные со случайными, но реалистичными значениями
- Записывает напрямую в базу данных

Минимальные требования:

- Не менее 3 полей в записи (помимо id и timestamp)
- Данные должны иметь смысл (не просто random числа)

2. База данных – PostgreSQL/MySQL(Percona, Maria)

Что хранит:

- Все сгенерированные события
- Минимум одна таблица с осмысленной структурой

3. Redash – дашборды и визуализации

Что нужно сделать:

- Подключить Redash к PostgreSQL/MySQL
 - Создать минимум 3 визуализации (графики/таблицы)
 - Собрать их в один дашборд
-

Варианты предметных областей (выбрать одну)

Вариант

Описание

Примеры полей

1. Погодная станция	Датчик отправляет показания	температура, влажность, давление, скорость ветра
2. Интернет-магазин ин	Поток заказов	товар, категория, цена, количество, город
3. Фитнес-трекер	Данные с устройства	шаги, пульс, калории, тип активности
4. Игровая статистика	События в игре	игрок, действие, очки, уровень
5. Такси/доставка	Поездки/заказы	откуда, куда, расстояние, цена, рейтинг
6. Свой вариант	Согласовать с преподавателем	—

Требования к реализации

Обязательно:

- **Данные должны быть осмысленными**
- **Docker Compose** – вся система запускается командой `docker-compose up`
- **Git-репозиторий** – код на GitHub/GitLab с осмысленной историей коммитов (не один коммит!)
- **README.md** – инструкция по запуску и описание проекта, скриншоты работы redash с дашбордом
- Генератор работает автоматически при запуске контейнеров
- В Redash создан дашборд с 3+ визуализациями

Результат необходимо отправить на почту skripnikov.mv@dvgu.ru до **начала зачетной недели** для получения допуска к зачету. В случае, если к работе будут вопросы, до начала зачетной недели необходимо успеть их исправить.

Для получения зачета необходимо защитить отправленную работу в установленное время (5-10 минут).

Дашборд в Redash и генерацию/хранение данных в БД необходимо будет продемонстрировать в рамках защиты (помним про docker volumes).