

## Домашнее задание

Дисциплина	Python для инженерии данных
Тема	Тема 11. Kafka
Форма проверки	<b>Самопроверка. Студент выполняет задание и самостоятельно проверяет его.</b>
Имя преподавателя	Дмитрий Клабуков
Время выполнения	1 час
Цель задания	Научиться работать с продюсерами и консьюмерами. Научиться обрабатывать поток данных
Инструменты для выполнения ДЗ	Kafka, Docker Compose
Правила приема работы	Прикрепите ссылку в LMS на выполненное задание в Google colab или GitHub (если вы использовали Jupyter Notebook) <b>Важно:</b> убедитесь в том, что по ссылке есть доступ в Google colab (иногда в колабе нет доступа для другого логина).
Критерии оценки	<b>Задание считается выполненным, если:</b> <ul style="list-style-type: none"><li>- прикреплена ссылка на файл с выполненным заданием</li><li>- доступ к файлу открыт</li><li>- код дает правильный ответ к задаче</li></ul> <b>Задание не выполнено, если:</b> <ul style="list-style-type: none"><li>- файл с заданием не прикреплен или отсутствует доступ по ссылке</li><li>- код выдаёт ошибку или дает неправильный ответ</li></ul>
Дедлайн	<b>16.12.2024</b>

### Задание

1. Разверните локально Kafka кластер с помощью Docker Compose.
2. Создайте топик `example_topic`.
3. Сгенерируйте данные, которые будут описывать действия клиента маркетплейса с его возможной активностью.
4. Загрузите данные в созданный топик.
5. Создайте консьюмер для чтения данных из полученного потока данных и проанализируйте данные:  
выведите на экран пользователей с наибольшим количеством действий (click, purchase).

### Пример формата данных:

```
{"user_id": 1, "action": "login", "timestamp": "2024-12-08T10:00:00"}
```

**Пример действий клиентов:**

login, logout, purchase, click

**Чек-лист самопроверки**

Критерии выполнения задания	Отметка о выполнении
Установлен jupyter notebook либо используется google colab	
Создан профиль на <a href="https://github.com">https://github.com</a> (при использовании jupyter notebook)	
Создан топик example_topic	
Данные, которые описывают действия клиента, сгенерированы и загружены в example_topic	
Создан консьюмер для чтения данных из потока. Код выводит на экран пользователей с наибольшим количеством действий (click, purchase)	
Прикреплена на учебной платформе ссылка на выполненное задание в Google colab или Github (если вы использовали jupyter notebook)	
Если используется Google colab, то по ссылке есть доступ (иногда в колабе нет доступа для другого логина)	