# Тестовое задание Data Engineer

## Описание источника

Имеется некоторый источник, который на ежедневной основе предоставляет данные в формате csv (см. файл cf\_test\_dataset.csv). В таблице указаны факты оплаты пользователями услуг в разрезе серверов и периодов действия услуги.

Описание полей файла:

user\_id – id клиента

server\_order\_id – id заказа

service\_id – id услуги

server\_configuration – конфигурация сервера

service\_start\_date – дата начала предоставления услуги (дата списания денежных средств)

service\_end\_date - дата окончания предоставления услуги

user\_id - id клиента

user\_name – имя клиента

user\_surname – фамилия клиента

price – стоимость заказа

## Задание

1. На основе прилагаемого файла (cf\_test\_dataset.csv) реализовать нормализованную структуру базы данных (MS SQL Server);
2. При помощи python-скриптов реализовать ETL-процесс доставки данных из файла до конечных таблиц;
3. Предусмотреть минимальное логирование результатов выполнения ETL-процесса
4. С помощью представлений и/или хранимых процедур построить следующие «отчеты»:

* Количество новых заказов, сделанных новыми клиентами помесячно. Клиент считается новым в месяц, когда начал потреблять услуги компании.
* Количество новых заказов серверов по каждому клиенту, в динамике помесячно в соответствии с примером ниже.

|  |  |  |  |
| --- | --- | --- | --- |
| Клиент | Месяц1 | Месяц2 | Месяц3 |
| Клиент1 | Кол-во заказов |  |  |
| Клиент2 | Кол-во заказов |  |  |

1. Результат выполнения тестового задания разместить на открытом репозитории