МИНИСТЕРСТВО НАУКИ и ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ

УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ

«Курганский государственный университет»

(ФГБОУ ВО «КГУ»)

Кафедра «Безопасность информационных и автоматизированных систем»

ОТЧЕТ

По лабораторной работе №1

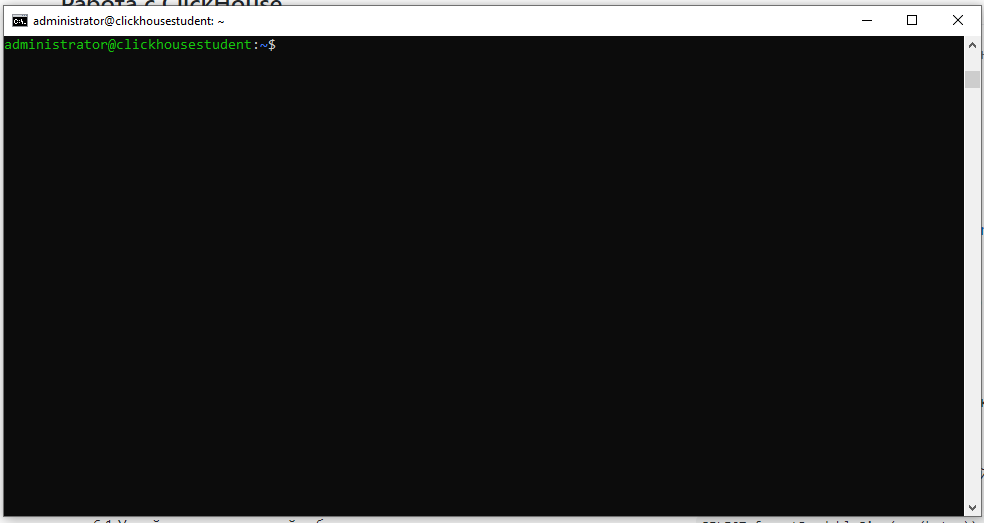
Работу выполнил

студент группы ИТ-40917

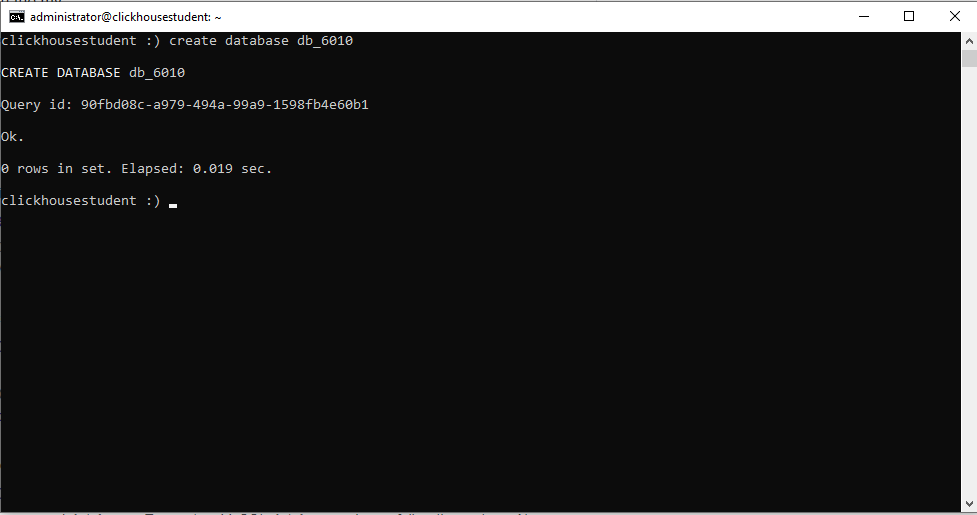
Скутин Роман

Курган 2020

1.Подключиться к лабораторной ВМ по SSH

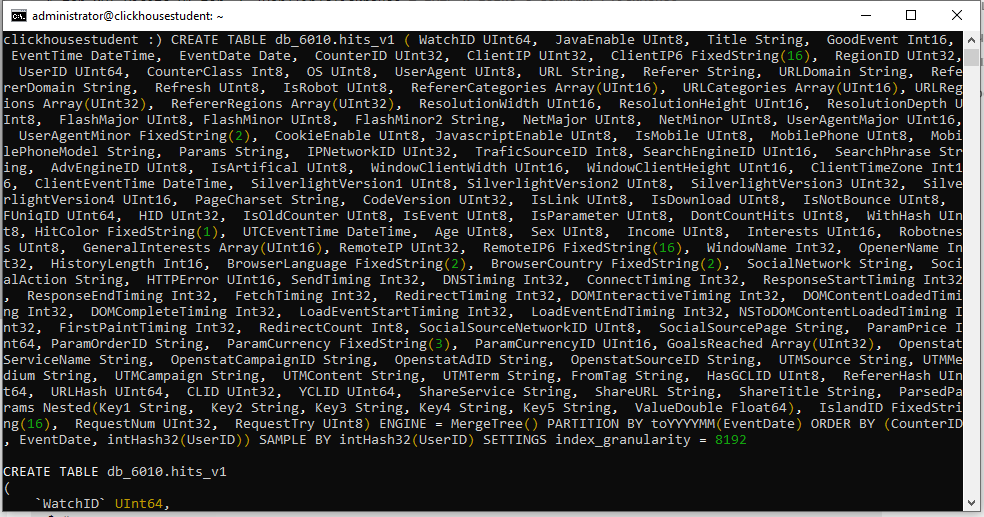


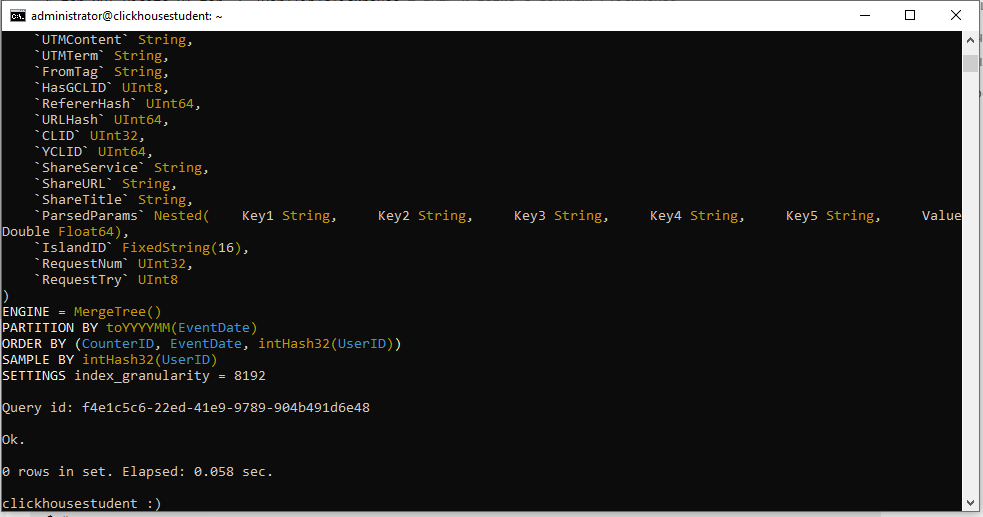
2.Создать БД с названием содержащим код своей группы



3.На диске находится файл hits\_v1.tsv требуется загрузить данные из него в таблицу

Создаём таблицу

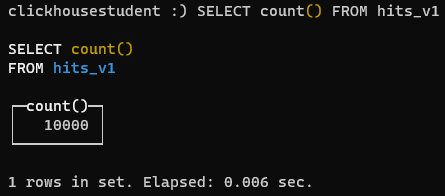




Загружаем строки с начала файла hits\_v1.tsv

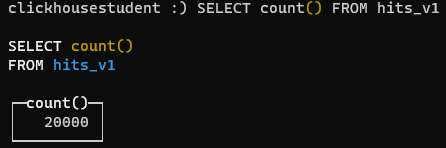
head --lines=10000 hits\_v1.tsv | clickhouse-client --password \*\*\*\*\*\*\*\*\*\*\* --query "INSERT INTO db\_6010.hits\_v1 FORMAT TSV" --max\_insert\_block\_size=100000

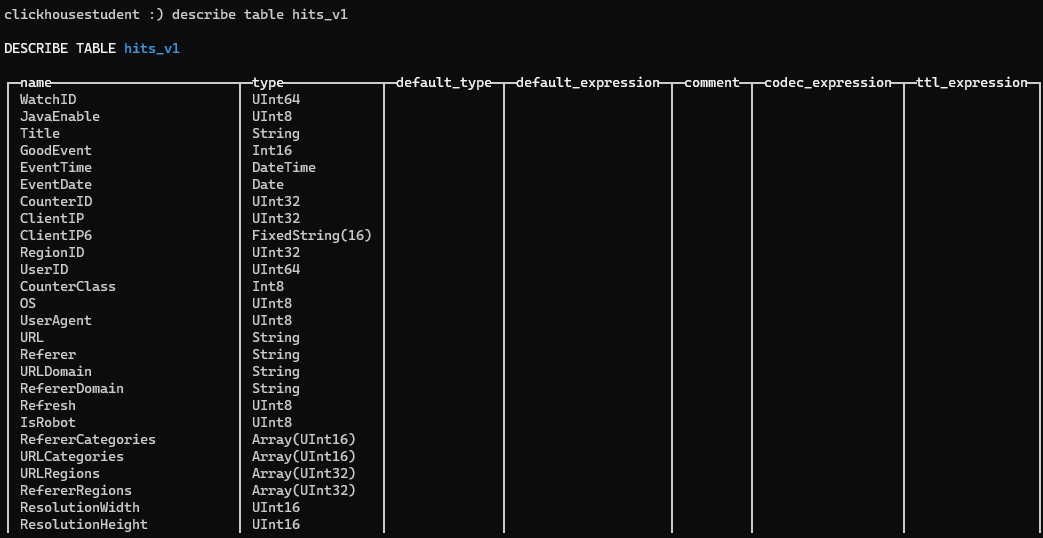
После загрузки строк в таблицу командой head



Загружаем строки с конца файла hits\_v1.tsv

tail -n 10000 hits\_v1.tsv | clickhouse-client --password \*\*\*\*\*\*\*\*\*\*\* --query "INSERT INTO db\_6010.hits\_v1 FORMAT TSV" --max\_insert\_block\_size=100000





Структура таблицы полностью соответствует тому, что было указано в запросе на создание

SELECT formatReadableSize(sum(bytes)) AS size, sum(rows) AS rows FROM system.parts WHERE active and database='db\_6010' and table='hits\_v1'



Число строк соответствует количеству загруженных

Вставим еще 10000 строк, пропустив первые 10000, скомбинировав команды head и tail

head -n 20000 hits\_v1.tsv | tail -n 10000 | clickhouse-client --password='\*\*\*\*\*\*\*\*\*\*\*' --query "insert into db\_6010.hits\_v1 format TSV" --max\_insert\_block\_size=10000



Объем данных возрос на 1.19 мегабайта

head -n 20000 hits\_v1.tsv | tail -n 10000 > temp.txt

Узнаем объем получившегося файла с помощью комныды ls -l --block-size=M.

Получился объём 10 мегабайт, что говорит о более компактном хранении данных в кликхаусе.