

Программа курса

«CHCD: Подготовка к сертификации ClickHouse Certified Developer»

О курсе: 1-дневный интенсивный курс для подготовки к сдаче сертификационного экзамена ClickHouse Certified Developer позволит систематизировать необходимые знания и навыки для успешной сдачи сертификационного экзамена ClickHouse Certified Developer. Курс построен на отработке практических примеров необходимых для выполнения заданий в экзаменационном практикуме.

Аудитория: Желающие подготовиться и сдать сертификационный экзамен (стоимость экзамена включена).

Уровень подготовки:

- опыт работы с SQL от 1 года
- английский язык (техн.)
- опыт работы с ClickHouse от 1 года
или **прослушанный курс «CLICH: Построение DWH на ClickHouse»**
- или **сдача предварительного экзамена**

Продолжительность курса: 8 академических часов, 1 день дистанционно

Тестирование: 2 часовой практический тест на английском языке доступный в любое время после прохождения курса в течении года.

Содержание программы

1. Сертификация ClickHouse Certified Developer

- Основные сведения
- Процедура и условия
- Домены

2. Знакомство с ClickHouse

3. Архитектура ClickHouse

4. Проектирование данных

- Создание новой базы данных
- Создание новой таблицы, удовлетворяющей заданным критериям или соответствующей определённому формату файла
- Выбор оптимальных типов данных для столбцов в зависимости от характера данных
- Определение эффективного ключа сортировки (primary key) на основе типов предполагаемых запросов в таблице MergeTree
- Определение и использование внешнего словаря (Dictionary)

5. Вставка и загрузка данных

- Загрузка локального файла в таблицу
- Загрузка файла из облачного хранилища (например, S3) в таблицу
- Загрузка файлов в формате Parquet, CSV или TSV
- Применение простых преобразований к столбцам во время вставки данных
- Вставка данных из одной таблицы в другую

6. Анализ данных

- Написание запроса, соответствующего заданным условиям

- Использование обычных функций в запросах: например, поиск подстрок в строковых столбцах или преобразование временной метки к началу интервала
- Использование агрегатных функций: например, нахождение max/min/sum/avg, количества уникальных значений или заданного квантиля
- Применение GROUP BY для группировки и агрегации значений по заданному временному интервалу или другим критериям

7. Материализованные представления

- Создание материализованного представления, сохраняющего результат неагрегирующего запроса

8. Агрегации в материализованных представлениях

- Создание материализованного представления, сохраняющего агрегированные данные с использованием движков AggregatingMergeTree или SummingMergeTree

9. Шардирование и репликация

- Работа с системными таблицами

10. JOIN операции

- Использование JOIN

11. Удаление и обновление данных

- Выполнение лёгкой (неблокирующей) операции удаления данных
- Реализация стратегии UPSERT с помощью движка ReplacingMergeTree
- Эффективное выполнение частых обновлений с использованием движка CollapsingMergeTree

12. Управление данными

- Работа с системными таблицами
- Создание и изменение таблиц

13. Оптимизация запросов

- Создание проекции (projection) в таблице для ускорения выполнения типовых запросов
- Определение skipping-индекса set или minmax для эффективного фильтра по диапазонам

14. Итоговое тестирование

Тестирование «ClickHouse Certified Developer» (**доступно только на английском языке**) будет доступно сразу по окончании курса после обязательной регистрации участника в течении 1 года (после регистрации).