

Лабораторная работа по теме “Выбор хранилища данных”

Задача:

Вы являетесь аналитиком данных в небольшой компании, которая решает перейти к современной системе управления данными. Вам поручено выбрать и обосновать выбор хранилища данных для компании. Ваша компания работает в сфере e-commerce и продает товары через онлайн-платформу. У вас есть следующие требования:

Объем данных: Ваша компания собирает 5 миллионов записей о продажах в месяц, 10 миллионов клиентов и 25 миллионов артикулов.

Аналитика: Вам необходимо обеспечить возможность проведения сложных аналитических запросов к данным для выявления тенденций продаж, сегментации клиентов и определения эффективности маркетинговых кампаний.

Масштабируемость: Ваша компания растет, и вам нужна система, которая легко масштабируется с ростом объема данных и нагрузки.

Реальное время: Вам нужна возможность анализа данных в реальном времени для мониторинга заказов и складской информации.

Задачи:

1. Выберите одно или несколько хранилищ данных (реляционная база данных, NoSQL, колоночное хранилище и т. д.), которые, по вашему мнению, наилучшим образом соответствуют требованиям компании.
2. Обоснуйте свой выбор, предоставив аргументы, почему данное хранилище данных является оптимальным для конкретных требований.
3. Опишите структуру данных, которые будут храниться в выбранном хранилище (таблицы, колонки и связи).
4. Предложите схему интеграции данных из разных источников в хранилище.
5. Опишите, как бы вы настроили систему для обработки данных в реальном времени.
6. Опишите команду, которая должна реализовывать предложенную Вами схему.
7. Опишите команду, которая должна поддерживать предложенную Вами схему.

Результат должен быть предоставлен в виде *.docx или *.pptx файла со схемами

Deadline: 16 сентября