

Лекция 1: Практическое задание

Курс: Разработчик ХД



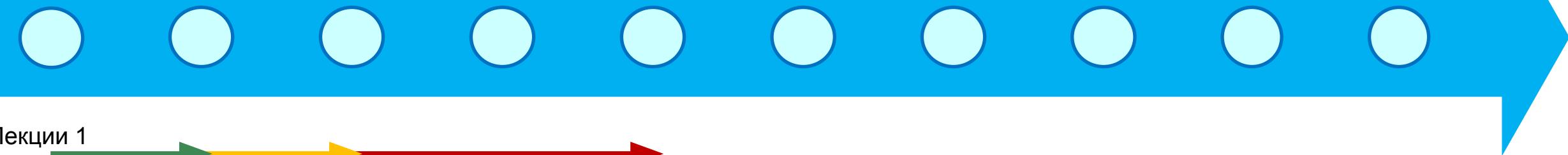
Copyright © 2018–2025 by DataTech. All Rights Reserved.





График лекций и сдачи практических заданий

| Лекция 1 | Лекция 2 | Лекция 3 | Лекция 4 | Лекция 5 | Лекция 6 | Лекция 7 | Лекция 8 | Крайний срок сдачи всех ДЗ | Дата проверки последнего ДЗ |
|----------|----------|----------|----------|----------|----------|----------|----------|----------------------------|-----------------------------|
| 16.09 | 19.09 | 23.09 | 26.09 | 30.09 | 03.10 | 07.10 | 10.10 | 14.10 | 16.10 |



ДЗ Лекции 1

ДЗ Лекции 2

ДЗ Лекции 3

ДЗ Лекции 4

ДЗ Лекции 5

ДЗ Лекции 6

ДЗ Лекции 7

ДЗ Лекции 8

Максимум 10 баллов

Максимум 6 баллов

Максимум 2 балла



Лекция 1. Практическое задание (1 из 2)

1. Сравните базу данных и хранилище данных по следующим критериям. Добавьте **собственные** критерии сравнения к перечисленным, если считаете их важными.
 - Решаемые бизнес-задачи.
 - Глубина хранения данных: актуальные (дни и месяцы), исторические (годы), те и другие.
 - Приоритетный тип **нагрузки**: транзакционный (OLTP), аналитический (OLAP).
 - **Затраты** на хранение данных: низкие, средние, высокие.
 - Тип **модели** данных (опционально, если знаете).
 - Пример использования **базы** данных.
 - Пример использования **хранилища** данных.
2. Сформулируйте **ключевое отличие(я)** между хранилищем данных (Data Warehouse) и озером данных (Data Lake) с точки зрения работы с данными конечных пользователей, как вы это понимаете.



Лекция 1. Практическое задание (2 из 2)

3. Сравните SMP-СУБД, МРР-СУБД и распределенную СУБД (Hadoop) по следующим критериям.

Добавьте собственные критерии сравнения к перечисленным, если считаете их важными.

- Тип **архитектуры**: разделяемые ресурсы (Shared Everything), частично-разделяемые ресурсы (Partially Shared), неразделяемые ресурсы (Shared Nothing), другая.
- Возможность **вертикального** масштабирования: да, нет, ограниченная.
- Возможность **горизонтального** масштабирования: да, нет, ограниченная.
- Примерные **объемы** постоянно хранимых данных.
- Основные **типы** хранимых данных (возможно несколько вариантов): структурированные, слабоструктурированные, неструктурированные.
- **Стоимость** владения из расчета в 1 терабайт постоянных данных: низкая, средняя, высокая.
- Производительность в **транзакционной** (OLTP) нагрузке: низкая, средняя, высокая.
- Производительность в **аналитической** (OLAP) нагрузке: низкая, средняя, высокая.

4. Предложите свой бизнес-сценарий, где лучшим выбором будет **резидентная** (in-memory) СУБД.

- Результаты практического задания лекции 1 принимаются в файлах форматов .doc, .docx, .txt.