Вопросы

к экзамену по дисциплине

«Методы сбора, хранения, обработки и анализа данных»

1. Логическое проектирование систем хранения данных. Сбор информации. Документация. Минимальные информационные требования. Источники для выявления правил данных.
2. Логическое проектирование систем хранения данных. Реляционная и нереляционная модели. Сущности атрибуты, отношения и бизнес-правила. Методологии моделирования. Нормализация и денормализация. План проекта использования данных.
3. Физическое проектирование и реализация. Проблемы использования данных, связанные с размером данных. Влияние физических характеристик.
4. Физическое проектирование и реализация. Основные топологии приложений. Схема физического проектирования. Обеспечение целостности данных. Расширенный доступ к данным.
5. Физическое проектирование и реализация. Обеспечение безопасности. Определение требований к аппаратным средствам. Определение параметров роста данных. Определение параметров архивирования.
6. Иерархии. Виды организации иерархических данных. Иерархический тип данных в SQL Server.
7. Иерархии. Хранение иерархических данных. Иерархические запросы в Oracle.
8. Обобщенные табличные выражения. Применение ОТВ. Рекурсивные ОТВ.
9. Графы. Графовые базы данных. Ключевое слово MATCH.
10. Расширенные группировки. Применение расширенных группировок в SQL Server и в Oracle.
11. Аналитические функции. Виды аналитических функций. Синтаксис аналитических функций.
12. Преобразование данных в SQL. Применение оператора MERGE. Применение операторов PIVOT и UNPIVOT.
13. Объектные типы данных в SQL Server. Применение объектных типов данных. Объектные типы данных в SQL Server.
14. Сборки. Применение сборок в SQL Server. Виды программных конструкций, используемых в сборках. Регистрация сборок. Использование сборок.
15. Пространственные данные. ГИС-приложения. Импорт пространственных данных из ГИС-систем. Виды пространственных данных.
16. Пространственные данные. SRID. Методы и свойства пространственных данных. Индексирование пространственных данных.
17. Объектные типы данных в Oracle. Атрибуты. Методы. Объектные таблицы. Объектные представления.
18. Объектные типы данных в Oracle. Индексирование объектных данных.
19. Коллекции в Oracle. Виды коллекций. Методы и исключения коллекций. Множественная обработка записей c выражением FORALL.
20. Коллекции в Oracle. Сравнение коллекций. MULTISET.
21. Коллекции в Oracle. Множественная обработка записей c выражением BULK COLLECT.
22. Динамический SQL. Использование динамического SQL в SQL Server и в Oracle.
23. Программные конструкции PL/SQL. Табличные функции.
24. Программные конструкции PL/SQL. Конвейерные функции.
25. Программные конструкции PL/SQL. Пакеты. Пакетные переменные. Пакетные курсоры. Пакетные исключения.
26. Технологии высокой доступности.
27. Ретроспективные запросы. Настройка ретроспективных запросов.
28. Задания. Системные пакеты DBMS\_JOB и DBMS\_SCHEDULER в Oracle.
29. Резервное копирование и восстановление. Применение резервного копирования в SQL Server.
30. Резервное копирование и восстановление. Применение резервного копирования в Oracle.
31. Репликация. Участники репликации. Настройка репликации.
32. Репликация. Топологии репликации. Виды репликации.
33. Настройка объектов базы данных. Кэширование. Журналирование. Параллельный доступ.
34. Основы бизнес-аналитики. Хранилища данных и киоски данных. Проектирование хранилищ данных.
35. Основы бизнес-аналитики. Кубы и их архитектура. Агрегирование. Уровень агрегирования. Физическое хранение куба.
36. Терминология служб SSAS. Разработка и просмотр многомерного куба.
37. NoSQL решения. MongoDB. Коллекции. Документы. Основные операции с документами и коллекциями.
38. MongoDB. Поиск в документах. Проекции. Ограничение выборки. Сортировка.
39. MongoDB. Индексирование. Виды индексов. Планы запросов. Оценка планов. Хинты.
40. MongoDB. Индексирование. Пространственные индексы. Методы поиска пространственных данных с учетом индекса.
41. MongoDB. Индексирование. Текстовый индекс. Различные виды текстовых индексов. Индексы с весами. Стемминг. Релевантность.
42. MongoDB. Фреймворк агрегации. Назначение. Пайплайн. Стадии.
43. MongoDB. Аутентификация. Роли. Пользователи. Привилегии. Встроенные роли.
44. MongoDB. Устройство базы данных. Файлы. Механизмы чтения и записи. Ограничения целостности. Типы данных.
45. Анализ данных в Python. Специализированные библиотеки.
46. Анализ данных в Python. Основные понятия статистики. Проверка гипотез. Статистический вывод.
47. Анализ данных в Python. Сравнение средних значений.
48. Анализ данных в Python. Корреляция и регрессия.
49. Анализ данных в Python. Применение метода Bootstrap для алгоритма деревьев решений.
50. Анализ данных в Python. Применение методов Airflow.