

УТВЕРЖДЕНА

Решением Ученого совета
АНО ВО «Центральный университет»
«07» марта 2024 г.
Протокол №1

**Рабочая программа дисциплины (модуля)
«Базы данных: Хранилища»**

Направление подготовки: 38.03.05 Бизнес-информатика

Направленность (профиль) подготовки: Бизнес-аналитика

Квалификация (степень) выпускника: бакалавр

Форма обучения: очная

Срок освоения программы: 4 года

Год набора: 2024

**Москва
2024**

Содержание

1. Краткая характеристика дисциплины (модуля)	3
2. Перечень планируемых результатов обучения	4
3. Тематический план	6
4. Содержание дисциплины (модуля)	7
5. Учебно-методическое обеспечение	8
6. Материально-техническое обеспечение	8
7. Методические и оценочные материалы	10

1. Краткая характеристика дисциплины (модуля)

Рабочая программа дисциплины (модуля) «Базы данных: Хранилища» составлена в соответствии с федеральным государственным образовательным стандартом высшего образования – бакалавриат по специальности 38.03.05 Бизнес-информатика, профиль Бизнес-аналитика, утвержденный приказом Министерства науки и высшего образования Российской Федерации № 838 от 29.07.2020 года.

Изучение дисциплины (модуля) «Базы данных: Хранилища» формирует ключевые навыки работы с системами хранения и анализа больших данных, что критически важно для поддержки бизнес-аналитики и принятия управленческих решений. Она обеспечивает понимание современных технологий обработки данных, включая ETL-процессы, OLAP-системы и облачные хранилища.

Место дисциплины (модуля) в структуре образовательной программы

Настоящая дисциплина (модуль) включена в учебный план по программе подготовки бакалавриата по направлению 38.03.05 Бизнес-информатика, профиль Бизнес-аналитика и входит в вариативную часть Блока 1, формируемую участниками образовательных отношений.

Дисциплина (модуль) является выборной и доступна для изучения на 3 или 4 курсе в 5, 6, 7, 8 семестрах на выбор.

Цель изучения дисциплины (модуля): освоение принципов проектирования, оптимизации и эксплуатации хранилищ данных для эффективного управления информацией и её анализа в корпоративных системах.

Задачи изучения дисциплины (модуля) направлены на формирование у студентов следующий знаний, умений и навыков:

- знание концепций хранения;
- знание принципов работы хранилищ данных;
- знание способов обработки больших данных;
- знание разновидностей хранилищ;
- знание актуального набора инструментов;
- знание способов применения концепций хранения и обработки данных к рабочим задачам;
- знание методов адаптации стандартных моделей данных к рабочей задаче;
- умение выбирать подходящий набор инструментов под задачу;
- умение проектировать хранилища данных;
- умение проектировать модели данных;
- умение оптимизировать количество используемых ресурсов.

2. Перечень планируемых результатов обучения

Компетенции, формируемые в результате освоения дисциплины (модуля) при проведении учебных занятий в форме контактной работы обучающихся с педагогическими работниками Университета и в форме самостоятельной работы обучающихся:

Компетенция	Содержание компетенции	Индикатор компетенции	Перечень планируемых результатов обучения по дисциплине (модулю)
УК-1.	Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач	УК-1.1.	Знает методы поиска и анализа информации в области аналитики, основные принципы критической оценки источников информации и их релевантности.
		УК-1.2.	Умеет критически оценивать источники информации и синтезировать данные из различных источников для решения задач, применять системный подход к анализу и решению комплексных проблем
		УК-1.3.	Имеет практический опыт работы с современными инструментами и технологиями для обработки информации, формулировании и структурировании задач на основе полученной информации
УК-2.	Способен определять круг задач в рамках поставленной цели и выбирать оптимальные способы их решения, исходя из действующих правовых норм, имеющихся ресурсов и ограничений	УК-2.1.	Знает действующие правовые нормы, регулирующие деятельность в области решения задач, основные методы и подходы к определению круга задач
		УК-2.2.	Умеет определять круг задач в рамках поставленной цели, выбирать оптимальные способы решения задач, учитывая имеющиеся ресурсы и ограничения
		УК-2.3.	Имеет практический опыт применения знаний о правовых нормах и ресурсах в реальных ситуациях, разработки и реализации решений в соответствии с установленными ограничениями
ОПК-2.	Способен проводить исследование и анализ рынка информационных систем и информационно-коммуникационных технологий, выбирать рациональные решения для управления бизнесом	ОПК-2.1.	Знает основные тенденции и характеристики рынка информационных систем и информационно-коммуникационных технологий
		ОПК-2.2.	Умеет проводить исследование и анализ рыночной информации для оценки потребностей бизнеса и выбора оптимальных

			решений
		ОПК-2.3.	Имеет практический опыт в разработке и внедрении стратегий управления бизнесом на основе анализа рынка информационных технологий
ПК-2.	Способен использовать соответствующий математический аппарат и инструментальные средства для обработки, анализа и систематизации информации по теме исследования для решения задач профессиональной деятельности	ПК-2.1.	Знает основные математические методы и инструментальные средства, применяемые для обработки и анализа информации
		ПК-2.2.	Умеет эффективно использовать математический аппарат для систематизации данных и решения профессиональных задач
		ПК-2.3.	Имеет практический опыт работы с инструментами анализа информации в рамках исследовательских проектов
ПК-3.	Способен готовить научно-технические отчеты, презентации, научные публикации по результатам выполненных исследований	ПК-3.1.	Знает требования и стандарты оформления научно-технических отчетов, презентаций и публикаций
		ПК-3.2.	Умеет структурировать и представлять результаты исследований в ясной и доступной форме
		ПК-3.3.	Имеет практический опыт подготовки и публикации научных материалов, отражающих результаты выполненных исследований
ПК-8.	Способен под руководством специалиста более высокой категории осуществлять планирование и организацию проектной деятельности на основе стандартов управления проектами	ПК-8.1.	Знает принципы и стандарты управления проектами
		ПК-8.2.	Умеет разрабатывать планы и организовывать проектную деятельность в соответствии с установленными стандартами
		ПК-8.3.	Имеет практический опыт участия в проектной работе, включая планирование и координацию задач

3. Тематический план

№п/п	Наименование раздела дисциплины (модуля)	Трудоемкость, академические часы				ТКУ (текущий контроль успеваемости)
		<i>Очная форма</i>				
		Контактная работа		Контроль	Самостоятельная работа	
Лекции	Семинары (практические занятия)					
1	Введение	3	3		16	Подготовка к семинару, Домашние задания
2	Концепции хранения и обработки больших данных	3	3		16	Подготовка к семинару, Домашние задания
3	Принципы работы хранилищ данных	4	4		16	Подготовка к семинару, Домашние задания
4	Виды хранилищ данных	4	4		16	Подготовка к семинару, Домашние задания, Контрольная работа
5	Модели хранения данных	3	3		16	Подготовка к семинару, Домашние задания
6	Инструменты и среды разработки	3	3		16	Подготовка к семинару, Домашние задания
7	Оптимизация и производительность	4	4		16	Подготовка к семинару, Домашние задания
8	Проектирование хранилищ данных	4	4		12	Подготовка к семинару, Домашние задания
	<i>Экзамен</i>			6		
	Итого:	30	30	6	124	
	Объем дисциплины (модуля) (в ак. ч.)	190				
	Объем дисциплины (модуля) (в зач. ед.)	5				

4. Содержание дисциплины (модуля)

№п/п	Наименование раздела дисциплины (модуля)	Содержание дисциплины (модуля) по темам
1	Введение	Основные понятия хранилищ данных. Роль хранилищ в аналитике. Эволюция систем хранения. Бизнес-ценность хранилищ. Современные требования к данным.
2	Концепции хранения и обработки больших данных	Принципы распределенного хранения. Обработка структурированных и неструктурированных данных. Масштабируемость систем. Обеспечение отказоустойчивости. Репликация и шардинг данных.
3	Принципы работы хранилищ данных	ETL-процессы (извлечение, преобразование, загрузка). Разделение транзакционных и аналитических систем. Обеспечение целостности данных. Механизмы обновления информации. Поддержка исторических данных.
4	Виды хранилищ данных	OLTP vs OLAP системы. Корпоративные хранилища (EDW). Витрины данных (Data Marts). Озера данных (Data Lakes). Облачные хранилища.
5	Модели хранения данных	Реляционная модель. Звездообразная схема. Снежинка (Snowflake). Колоночное хранение. Графовые базы данных.
6	Инструменты и среды разработки	SQL-серверы (Oracle, MS SQL, PostgreSQL). BI-платформы (Power BI, Tableau). Облачные решения (AWS Redshift, Google BigQuery). Средства ETL (Informatica, Talend). Фреймворки для больших данных (Hadoop, Spark).
7	Оптимизация и производительность	Индексирование данных. Кэширование запросов. Партиционирование таблиц. Оптимизация запросов. Балансировка нагрузки.
8	Проектирование хранилищ данных	Методологии проектирования (Kimball, Inmon). Определение фактов и измерений. Нормализация и денормализация. Обеспечение безопасности данных. Документирование архитектуры.

5. Учебно-методическое обеспечение

Университет располагает полным набором лицензионного и свободно распространяемого программного обеспечения, включая продукты отечественного производства.

Каждый студент в течение всего периода обучения получает индивидуальный неограниченный доступ к электронно-библиотечной системе и электронной информационно-образовательной среде университета. Эти системы предоставляют возможность доступа к ресурсам из любой точки, где есть подключение к сети Интернет, как на территории университета, так и за его пределами.

Студентам обеспечен удаленный доступ к современным профессиональным базам данных и информационным справочным системам.

Основная литература:

1. Советов, Б. Я. Базы данных : учебник для вузов / Б. Я. Советов, В. В. Цехановский, В. Д. Чертовской. — 4-е изд., перераб. и доп. — Москва : Издательство Юрайт, 2025. — 403 с. — (Высшее образование). — ISBN 978-5-534-18479-2. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/559898>.

Дополнительная литература:

1. Парфенов, Ю. П. Постреляционные хранилища данных : учебное пособие для вузов / Ю. П. Парфенов ; под научной редакцией Н. В. Папуловской. — Москва : Издательство Юрайт, 2025. — 97 с. — (Высшее образование). — ISBN 978-5-534-21173-3. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/559502>.

6. Материально-техническое обеспечение

Университет располагает материально-технической базой, соответствующей действующим противопожарным правилам и нормам и обеспечивающей проведение всех видов дисциплинарной и междисциплинарной подготовки, практической и научно-исследовательской работ обучающихся, предусмотренных учебным планом.

Помещения, которые представляют собой учебные аудитории для проведения занятий лекционного типа, занятий семинарского (практического) типа, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, а также помещения для самостоятельной работы и помещения для хранения и профилактического обслуживания учебного оборудования. Помещения укомплектованы специализированной мебелью и техническими средствами обучения, служащими для представления учебной информации большой аудитории.

Изучение дисциплины (модуля) обеспечивается в учебных аудиториях, оснащенных:

- столами и стульями;
- компьютерной техникой;
- специализированным оборудованием, включая демонстрационное оборудование.

Помещения для самостоятельной работы обучающихся, в том числе приспособленные для использования инвалидами и лицами с ограниченными возможностями здоровья, оснащены компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа к в электронную информационно-образовательную среду Университета.

Обучающимся предоставляется доступ (в том числе удаленный) к ресурсам информационно-телекоммуникационной сети «Интернет», электронным ресурсам (в том числе электронным библиотечным системам, современным профессиональным базам данных и информационным справочным системам):

№	Наименование портала (издания, курса, документа)	Ссылка
1.	Научная электронная библиотека elibrary.ru библиотека	https://elibrary.ru/defaultx.asp
2.	База данных для IT-специалистов	https://habr.com
3.	База данных ScienceDirect	https://www.sciencedirect.com
4.	Официальный сайт Министерства науки и высшего образования Российской Федерации	https://minobrnauki.gov.ru/
5.	Федеральный портал «Российское образование»	https://www.edu.ru/
6.	Информационная система "Единое окно доступа к образовательным ресурсам"	http://window.edu.ru/
7.	Единая коллекция цифровых образовательных ресурсов	http://school-collection.edu.ru/
8.	Федеральный центр информационно - образовательных ресурсов	http://fcior.edu.ru/

Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), в том числе комплект лицензионного программного обеспечения, современные профессиональные базы данных и информационные справочные системы:

Наименование ПО	Производство	Лицензионное / свободно распространяемое
Операционные системы:		
Microsoft Imagine (Windows Client, Server)	зарубежное	лицензионное
Браузеры:		
Яндекс.Браузер	отечественное	свободно распространяемое
Google Chrome	зарубежное	свободно распространяемое
Офисные приложения:		
Microsoft Imagine (Visio, OneNote)	зарубежное	лицензионное
TeXstudio	зарубежное	свободно распространяемое
Adobe Acrobat Reader	зарубежное	свободно распространяемое
Программное обеспечение для планирования и учета времени:		
Toggle app	зарубежное	свободно распространяемое
Системы управления проектами:		
Microsoft Imagine (Project)	зарубежное	лицензионное
Системы управления базами данных:		
Microsoft Imagine (SQL Server)	зарубежное	лицензионное
Системы резервного копирования (backup):		
Acronis Backup Advanced for HyperV	зарубежное	лицензионное
Справочно-правовые системы:		
КонсультантПлюс: справочно-правовая система	отечественное	лицензионное
Средства антивирусной защиты:		
Kaspersky Endpoint Security для бизнеса Стандартный Russian Edition	отечественное	лицензионное
Среды разработки:		
Visual Studio Code	зарубежное	свободно распространяемое
Bash (Unix shell)	зарубежное	свободно распространяемое
Anaconda	зарубежное	свободно распространяемое
Robotic Operating System	зарубежное	свободно распространяемое
CopelliaSim	зарубежное	свободно распространяемое
Google Colaboratory	зарубежное	свободно распространяемое
Пакеты программных средств и библиотек:		

AutoPsy	зарубежное	свободно распространяемое
Interactive Disassembler (IDA)	зарубежное	свободно распространяемое
Системы управления библиографической информацией:		
Zotero	зарубежное	свободно распространяемое
Сервисы и службы:		
Bind	зарубежное	свободно распространяемое
Docker	зарубежное	свободно распространяемое

7. Методические и оценочные материалы

Методические указания для обучающихся по освоению дисциплины (модуля)

В процессе изучения дисциплины (модуля) «Базы данных: Хранилища» в рамках текущего контроля успеваемости используются такие виды учебной работы, как лекции, семинары, контрольная работа, домашние задания, а также различные виды самостоятельной работы обучающихся по заданию преподавателя, направленные на развитие навыков профессиональной лексики, закрепление практических профессиональных компетенций, поощрение инициатив.

Лекция – систематическое, последовательное, монологическое изложение преподавателем учебного материала, как правило, теоретического характера.

В процессе лекций рекомендуется вести конспект лекций: кратко и схематично фиксировать основные идеи, выводы и обобщения лекции; выделять важные мысли, ключевые слова и термины. Необходимо отметить вопросы или материалы, которые вызывают затруднения, и попытаться найти ответы в рекомендованной литературе. Если разобраться в материале не удастся, следует сформулировать вопрос и задать его преподавателю на консультации или во время семинарского (практического) занятия.

Участие в семинаре (аудиторная работа) – активная работа студента на семинаре, его ответы на вопросы преподавателя и участие в дискуссии.

Для успешного участия в семинаре студентам рекомендуется заранее ознакомиться с темой обсуждения, прочитать необходимые материалы и подготовить вопросы. Важно активно слушать и вовлекаться в дискуссию, высказывая свои мнения и аргументируя их. При ответах на вопросы преподавателя стоит быть уверенным, четким и логичным, опираясь на изученный материал. Также полезно поддерживать диалог с однокурсниками, чтобы обогатить обсуждение и расширить свои знания.

Домашнее задание – набор задач по темам недели.

При работе над домашними заданиями важно внимательно ознакомиться с требованиями и сроками выполнения. Рекомендуется разбивать задания на этапы, чтобы избежать перегрузки и лучше усвоить материал. Использовать различные источники информации, включая учебники и онлайн-ресурсы, для более глубокого понимания темы.

Контрольная работа – письменная работа с набором задач, которые нужно решить за ограниченное время.

Цель контрольной работы - получить специальные знания по одной или нескольким темам дисциплины (модуля) и продемонстрировать навыки их практического применения.

Самостоятельная работа – работа студентов, направленная на углубленное изучение отдельных тем и вопросов учебной дисциплины (модуля).

В процессе самостоятельной работы студенты взаимодействуют с рекомендованными материалами при минимальном участии преподавателя. Задачи студента включают работу с конспектами лекций (обработка текста), повторное изучение учебных материалов, планов и тезисов ответов, изучение дополнительных тем, выполнение учебно-исследовательских

заданий и другое.

Система оценивания результатов обучения по дисциплине (модулю)

Критерии получения уровня и оценивания сформированности компетенций по дисциплине (модулю) «Базы данных: Хранилища»

Оценивание уровня учебных достижений, обучающихся по дисциплине (модулю), осуществляется в виде текущего контроля успеваемости и промежуточной аттестации.

Промежуточная аттестация по дисциплине (модулю) осуществляется в форме **экзамена**, при этом проводится оценка компетенций, сформированных по дисциплине.

Для оценивания текущего контроля успеваемости и промежуточной аттестации используется десятибалльная шкала оценивания, которая соотносится с традиционной пятибалльной шкалой следующим образом:

Десятибалльная оценка	Пятибалльная оценка	Общая характеристика результата обучения по дисциплине (модулю)
10	Отлично	Студент полностью владеет знаниями, изложенными в рабочей программе, и глубоко осмысляет дисциплину. Он самостоятельно и логически последовательно отвечает на все вопросы, акцентируя внимание на наиболее важном. Умеет анализировать, сравнивать, классифицировать, обобщать, конкретизировать и систематизировать изученный материал, выделяя ключевые моменты и устанавливая причинно-следственные связи. Четко формулирует ответы, уверенно интерпретирует результаты анализов и других исследований, а также решает сложные задачи. Студент хорошо знаком с методами исследования, необходимыми для практической деятельности, и умеет связывать теоретические аспекты дисциплины (модуля) с практическими задачами.
9	Отлично	
8	Отлично	
7	Хорошо	Студент обладает знаниями предмета почти в полном объеме рабочей программы и самостоятельно, логически последовательно и всесторонне отвечает на все вопросы, акцентируя внимание на наиболее значимых моментах. Он умеет анализировать, сравнивать, классифицировать, обобщать, конкретизировать и систематизировать изученный материал, выделяя его ключевые аспекты и устанавливая причинно-следственные связи. Формулирует свои ответы, уверенно интерпретирует результаты анализов и других исследований, а также решает сложные ситуационные задачи. Студент хорошо знаком с методами исследования, необходимыми для практической деятельности, и умеет связывать теоретические аспекты предмета с практическими задачами.
6	Хорошо	
5	Удовлетворительно	Студент обладает базовыми знаниями по дисциплине (модулю), но испытывает трудности при самостоятельных ответах и использует неточные формулировки. В ходе ответов он допускает ошибки,
4	Удовлетворительно	

Десятибалльная оценка	Пятибалльная оценка	Общая характеристика результата обучения по дисциплине (модулю)
		касающиеся сути вопросов. Студент способен решать только самые простые задачи и владеет лишь минимальным набором методов исследования.
3	Не сдан	Студент не овладел обязательным минимумом знаний по предмету и не может ответить на вопросы, даже если преподаватель задает дополнительные наводящие вопросы.
2	Не сдан	
1	Не сдан	

Дисциплина (модуль) «Базы данных: Хранилища» оценивается следующим образом:

Активность	Вес	Количество	Описание
Домашние задания	30%	14	Набор задач по темам недели
Аудиторная работа	10%	14	Активная работа студента на семинаре
Контрольная работа	20%	1	Письменная работа с набором задач, которые нужно решить за ограниченное время
Экзамен	40%	1	Письменная или устная работа над заданием, направленным на проверку полученных знаний и навыков по дисциплине (модулю)

Формула расчёта итоговой оценки по дисциплине (модулю) «Базы данных: Хранилища»: « $0,3 \times$ среднее за домашние задания + $0,1 \times$ аудиторная работа + $0,2 \times$ контрольная работа + $0,4 \times$ экзамен».

Текущий контроль успеваемости обучающихся по дисциплине (модулю)

Примерные домашние задания

Домашнее задание по теме «Концепции хранения и обработки больших данных»

- Сравните технологии распределенного хранения (HDFS vs S3 vs Google Cloud Storage) – составьте таблицу с характеристиками.
- Проанализируйте кейс использования шардинга и репликации в крупной компании (Twitter, Uber или банковский сектор).
- Разработайте схему архитектуры для хранения и анализа логов веб-приложения (неструктурированные данные).
- Напишите эссе на тему "Проблемы отказоустойчивости в распределенных хранилищах: методы решения".
- Разверните кластер Apache Hadoop (или его эмуляцию) и загрузите тестовые данные.

Домашнее задание по теме «Принципы работы хранилищ данных»

- Разработайте ETL-пайплайн для извлечения и загрузки данных из CSV в SQL-хранилище (Python + SQL).
- Сравните методологии Kimball и Inmon – подготовьте презентацию с плюсами/минусами для разных бизнес-сценариев.
- Предложите стратегию инкрементальной загрузки данных для retail-аналитики.
- Ответьте на 10 тестовых вопросов по обеспечению целостности данных (первичные ключи, транзакции, ACID).
- Спроектируйте механизм хранения исторических данных для системы учета продаж.

Домашнее задание по теме «Инструменты и среды разработки»

1. Установите и настройте PostgreSQL + Pentaho для ETL – подготовьте инструкцию со скриншотами.
2. Сравните BI-инструменты (Power BI vs Tableau vs Metabase) – составьте таблицу с возможностями и ценами.
3. Загрузите датасет в Google BigQuery и выполните аналитический запрос.
4. Напишите скрипт на Python для автоматической выгрузки данных из API в хранилище (с использованием Airflow).
5. Протестируйте производительность – сравните скорость запросов в ClickHouse и MySQL.

Примерные вопросы для подготовки к семинарам

Вопросы к семинару по теме «Виды хранилищ данных»

1. Чем отличаются операционные базы данных (OLTP) и хранилища данных (OLAP)?
2. Какие преимущества дают колоночные хранилища (например, ClickHouse, Vertica) по сравнению со строковыми?
3. Как работают распределенные хранилища данных (HDFS, Amazon S3, Google Cloud Storage)?
4. В каких случаях используют NoSQL-хранилища (MongoDB, Cassandra) вместо реляционных?
5. Какие типы хранилищ лучше подходят для аналитики в реальном времени (real-time analytics)?

Вопросы к семинару по теме «Модели хранения данных»

1. В чем разница между схемами "звезда" (Star Schema) и "снежинка" (Snowflake Schema)?
2. Как выбирают уровень детализации (гранулярность) данных в хранилище?
3. Какие подходы к хранению исторических данных (SCD – Slowly Changing Dimensions) существуют?
4. Какова роль нормализации и денормализации в проектировании хранилищ?
5. Какие модели данных используют в Data Lake (структурированные, полуструктурированные, неструктурированные)?

Вопросы к семинару по теме «Оптимизация и производительность»

1. Какие методы индексации ускоряют запросы в хранилищах данных?
2. Как влияет партиционирование таблиц на производительность запросов?
3. Какие стратегии кэширования данных применяют в аналитических хранилищах?
4. Как можно оптимизировать ETL-процессы для уменьшения времени загрузки?
5. Какие метрики используют для оценки производительности хранилищ данных?

Примерные задания по контрольной работе

Контрольная работа

1. Дайте определение хранилища данных (Data Warehouse) и перечислите его ключевые характеристики.
2. Объясните, как эволюционировали системы хранения данных от файловых систем до современных облачных хранилищ.
3. Назовите три бизнес-задачи, которые решаются с помощью хранилищ данных, и приведите примеры.
4. Сравните требования к данным в операционных системах (OLTP) и аналитических хранилищах (OLAP).
5. Какие современные тренды в области хранения данных вы считаете наиболее

важными? Обоснуйте ответ.

6. Опишите принципы распределенного хранения данных (на примере HDFS или Amazon S3).

7. В чем разница между структурированными и неструктурированными данными? Приведите примеры форматов.

8. Какие методы масштабирования (горизонтальное vs вертикальное) применяются в хранилищах данных?

9. Объясните, как обеспечивается отказоустойчивость в распределенных системах (на примере репликации).

10. Что такое шардинг и в каких случаях он применяется? Приведите пример использования.

11. Опишите этапы ETL-процесса и приведите пример инструментов для каждого этапа.

12. Почему важно разделять транзакционные (OLTP) и аналитические (OLAP) системы?

13. Какие механизмы обеспечивают целостность данных в хранилищах?

14. Как организовано обновление данных в хранилище? Рассмотрите методы полной и инкрементальной загрузки.

15. Что такое Slowly Changing Dimensions (SCD) и какие типы SCD вы знаете?

16. Сравните OLTP и OLAP системы по критериям: структура данных, тип запросов, производительность.

17. Что такое корпоративное хранилище (EDW) и чем оно отличается от витрин данных (Data Marts)?

18. Опишите концепцию Data Lake и приведите примеры технологий для его реализации.

19. Какие преимущества дают облачные хранилища данных (на примере Google BigQuery, Snowflake)?

20. Приведите пример гибридного хранилища данных и объясните его архитектуру.

Задания для промежуточной аттестации по дисциплине (модулю)

№ п/п	Задание	Ответ	Компетенция
1.	Какой принцип отличает хранилище данных (DWH) от операционной базы данных (OLTP)? а) Хранилище оптимизировано для транзакций б) Хранилище содержит исторические данные и поддерживает аналитику в) Хранилище не требует ETL-процессов г) Хранилище работает только с неструктурированными данными	б	УК-1
2.	Какой механизм обеспечивает отказоустойчивость в распределенных хранилищах?	Репликация	УК-2
3.	Какое хранилище предназначено для хранения сырых данных любого формата? а) Data Mart б) EDW в) OLTP-система г) Data Lake	г	ОПК-2
4.	Какая модель хранения использует центральную таблицу фактов и связанные таблицы измерений? а) Реляционная б) Звездообразная схема в) Снежинка г) Графовая	б	ПК-2

5.	Какой инструмент НЕ относится к ETL-системам? а) Talend б) Informatica в) Power BI г) Apache NiFi	в	ПК-3
6.	Какая методология проектирования DWH предполагает построение единого централизованного хранилища? а) Inmon б) Kimball в) Data Vault г) Agile	а	ПК-8
7.	Как называется процесс извлечения, преобразования и загрузки данных?	ETL	УК-1
8.	Как называется метод разделения данных на логические части для ускорения запросов?	Шардинг	УК-2
9.	Как называются специализированные хранилища для отдельных подразделений компании?	Витрины данных (Data Marts)	ОПК-2
10.	Как называется нормализованная версия звездообразной схемы?	Снежинка (Snowflake)	ПК-2
11.	Как называется язык запросов для реляционных баз данных?	SQL	ПК-3
12.	Как называется процесс упрощения структуры таблиц для ускорения запросов?	Денормализация	ПК-8