ЦЕНТРАЛЬНЫЙ УНИВЕРСИТЕТ

УТВЕРЖДЕНА

Решением Ученого совета
АНО ВО «Центральный университет»
«25» мая 2026 г.
Протокол №2

Рабочая программа дисциплины (модуля)
«Хранилища больших данных»

Направление подготовки: *02.03.01 Математика и компьютерные науки*

Направленность (профиль) подготовки: *Дизайн и разработка информационных продуктов*

Квалификация (степень) выпускника: *бакалавр*

Форма обучения: *очная*

Срок освоения программы: *4 года*

Год набора: *2026*

ДОКУМЕНТ ПОДПИСАН
ЭЛЕКТРОННОЙ ПОДПИСЬЮ

Сертификат: 0215B0C000D1B3FEB44A37EE631E0C27D4
Владелец: Ивашкевич Евгений Васильевич
Действителен: с 13.01.2026 до 13.04.2027

Электронный документ

Москва
2026

Содержание

1. **Краткая характеристика дисциплины (модуля)** 3
2. **Перечень планируемых результатов обучения** 5
3. **Тематический план** 5
4. **Содержание дисциплины (модуля)** 9
5. **Учебно-методическое обеспечение** 10
6. **Материально-техническое обеспечение** 10
7. **Методические и оценочные материалы** 12

1. Краткая характеристика дисциплины (модуля)

Рабочая программа дисциплины (модуля) «Хранилища больших данных» составлена в соответствии с федеральным государственным образовательным стандартом высшего образования – бакалавриат по специальности 02.03.01 Математика и компьютерные науки, профиль Дизайн и разработка информационных продуктов, утвержденный приказом Министерства науки и высшего образования Российской Федерации № 807 от 23.08.2017 года.

Изучение дисциплины (модуля) «Хранилища больших данных» позволяет подготовить специалистов, способных строить надёжные и масштабируемые хранилища данных, обеспечивающие высокое качество информации и эффективную поддержку аналитики. Это ключевой элемент современной data-инфраструктуры, обеспечивающий доверие к данным и скорость принятия решений в бизнесе.

Место дисциплины (модуля) в структуре образовательной программы

Настоящая дисциплина (модуль) включена в учебный план по программе подготовки бакалавриата по направлению 02.03.01 Математика и компьютерные науки, профиль Дизайн и разработка информационных продуктов и входит в обязательную часть Блока 1.

Дисциплина (модуль) изучается на 4 курсе в 7 семестре.

Цель изучения дисциплины (модуля): сформировать у обучающихся системное понимание проектирования и эксплуатации хранилищ данных, включая моделирование, управление качеством, организацию загрузок и взаимодействие с бизнесом для поддержки аналитических процессов.

Задачи изучения дисциплины (модуля):

- изучить архитектуру и жизненный цикл выполнения Java-программ — от компиляции и структуры bytecode до загрузки классов, интерпретации и JIT-компиляции;
- освоить основные подходы к моделированию данных в хранилищах (звезда, снежинка, факты и измерения) для эффективной организации аналитических данных;
- изучить архитектуру слоёв хранилища (staging, core, marts), принципы SCD и реализации инкрементальных загрузок;
- научиться оценивать и обеспечивать качество данных, отслеживать lineage и управлять доступами в рамках данных governance;
- приобрести навыки проектирования витрин данных, ориентированных на конкретные аналитические сценарии и бизнес-вопросы;
- разработать мини-проект хранилища с документацией, контрактами данных и коммуникацией с заинтересованными сторонами.

В результате освоения дисциплины (модуля) обучающийся должен:

знать:

- подходы к моделированию данных в DWH: звезда/снежинка, факты/измерения;
- слои и принципы построения хранилища: staging/core/marts, SCD, инкрементальные загрузки;
- качество данных, lineage, governance и доступы;
- витрины и поддержка аналитических сценариев;

уметь:

- проектировать модель данных под бизнес-вопросы и запросы аналитики;
- планировать загрузки и обновления (инкременты, backfill, поздние данные);
- определять метрики качества данных и внедрять проверки;

владеть:

- навыками инженерного проектирования DWH: контракты данных, документация, стандарты;
- навыкам коммуникации с аналитиками и владельцами доменов для уточнения требований;
- навыками подготовки мини-проекта модели и витрины.

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 искусственного интеллекта, 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-1.	Способен консультировать и использовать фундаментальные знания в области математического анализа, комплексного и функционального анализа алгебры, аналитической геометрии, дифференциальной геометрии и топологии, дифференциальных уравнений, дискретной математики и математической логики, теории вероятностей, математической статистики и случайных процессов, численных методов, теоретической механики в профессиональной деятельности	ОПК-1.1.	Знает основные концепции и теории в области математического анализа и смежных дисциплин; методы и подходы, используемые в различных областях математики
		ОПК-1.2.	Умеет применять математические методы для решения профессиональных задач
		ОПК-1.3.	Имеет практический опыт разработки и реализация математических моделей в профессиональной деятельности
ОПК-5.	Способен понимать принципы работы современных информационных технологий и использовать их для решения задач профессиональной деятельности	ОПК-5.1.	Знает технологии, необходимые для прикладного программирования, включая современные функциональные языки программирования, а также основные принципы и понятия, применяемые при использовании

			компьютерных сетей
		ОПК-5.2.	Умеет пользоваться технологиями прикладного программирования, включая среды высокоуровневого программирования
		ОПК-5.3.	Имеет практический опыт использования технологий прикладного программирования
ПК-2.	Способен решать типовые задачи профессиональной деятельности, опираясь на информационную и библиографическую культуру, используя информационно-коммуникационные технологии и учитывая основные требования информационной безопасности	ПК-2.1.	Знает основы информационной и библиографической культуры, а также принципы информационной безопасности и применения информационно-коммуникационных технологий в профессиональной деятельности
		ПК-2.2.	Умеет эффективно использовать информационно-коммуникационные технологии для решения стандартных задач профессиональной деятельности, учитывая требования информационной безопасности
		ПК-2.3.	Имеет опыт работы с информационными ресурсами и технологиями в области профессиональной деятельности, включая соблюдение норм информационной безопасности
ПК-4.	Способен разрабатывать программное обеспечение для решения прикладных задач в области профессиональной	ПК-4.1.	Знает основные принципы разработки программного обеспечения и методы

	деятельности под руководством более опытного специалиста		решения прикладных задач в сфере профессиональной деятельности
		ПК-4.2.	Умеет разрабатывать программное обеспечение под руководством специалистов более высокой категории, используя современные технологии и инструменты
		ПК-4.3.	Имеет опыт участия в разработке программного обеспечения для решения прикладных задач в команде под руководством более опытных специалистов

3. Тематический план

№п/п	Наименование раздела дисциплины (модуля)	Трудоемкость, академические часы					ТКУ (текущий контроль успеваемости)
		Очная форма					
		Контактная работа			Контроль	Самостоятельная работа	
		Лекции	Семинары (практические занятия)	Консультации			
1	Введение в DWH	4	4	1		8	Домашние задания
2	Моделирование	4	4	1		8	Домашние задания
3	История и изменения	4	4	1		9	Домашние задания
4	Витрины	4	4	1		9	Домашние задания
5	Качество и governance	5	5	1		9	Домашние задания
6	Доступы и безопасность	5	5	1	2	9	Мини-проект
	Экзамен				2		
Итого:		***26***	***26***	***6***	***4***	***52***	
Объем дисциплины (модуля) (в ак. ч.)		***114***					
Объем дисциплины (модуля) (в зач. ед.)		***3***					

4. Содержание дисциплины (модуля)

№п/п	Наименование раздела дисциплины (модуля)	Содержание дисциплины (модуля) по темам
1	Введение в DWH	Зачем DWH, слои, типовые сценарии и анти-паттерны
2	Моделирование	Факты/измерения, гранулярность, звезда/снежинка
3	История и изменения	SCD, инкременты, поздние данные, backfill
4	Витрины	Проектирование marts, производительность запросов, агрегаты
5	Качество и governance	Проверки, lineage, документация, владение данными
6	Доступы и безопасность	Роли, маскирование, аудит, политики хранения

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. Карау, Х. Изучаем Spark. Молниеносный анализ данных : практическое руководство / Х. Карау, Э. Конвински, П. Венделл, З. Матей ; пер. с англ. - 2-е изд. - Москва : ДМК Пресс, 2023. - 305 с. - ISBN 978-5-89818-320-2. - Текст : электронный. - URL: https://znanium.ru/catalog/product/2102607.

2. Харенслак, Б. Apache Airflow и конвейеры обработки данных : практическое руководство / Б. Харенслак, Д. дэ Руйтер ; пер. с англ. Д. А. Беликова. - Москва : ДМК Пресс, 2022. - 502 с. - ISBN 978-5-97060-970-5. - Текст : электронный. - URL: https://znanium.ru/catalog/product/2155905.

3. Apache Kafka. Потоковая обработка и анализ данных : практическое руководство / Г. Шапира, Т. Палино, Р. Сиварам, К. Петти. - 2-е изд. - Санкт-Петербург : Питер, 2023. - 512 с. - (Серия «Бестселлеры O'Reilly»). - ISBN 978-5-4461-2288-2. - Текст : электронный. - URL: https://znanium.ru/catalog/product/2123357.

Дополнительная литература:

1. Карпова, И. П. Базы данных : учебное пособие / И. П. Карпова. - Санкт-Петербург : Питер, 2021. - 240 с. - (Серия «Учебное пособие»). - ISBN 978-5-4461-9681-4. - Текст : электронный. - URL: https://znanium.com/catalog/product/1857026.

2. Кэмпбелл, Л. Базы данных. Инжиниринг надежности : практическое руководство / Л. Кэмпбелл, Ч. Мейджорс. - Санкт-Петербург : Питер, 2020. - 304 с. - (Серия «Бестселлеры O'Reilly»). - ISBN 978-5-4461-1310-1. - Текст : электронный. - URL: https://znanium.com/catalog/product/1756135.

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;
- механическими калькуляторами;

— 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 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) «Хранилища больших данных» в рамках текущего контроля успеваемости используются такие виды учебной работы, как лекции, семинары, коллоквиумы, 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 удается, следует сформулировать вопрос и задать его преподавателю на консультации или во время семинарского (практического) занятия.

Семинар – это форма учебной деятельности, проводимая в учебном заведении под руководством преподавателя, где студенты активно участвуют в обсуждениях, практических заданиях и других формах взаимодействия.

Для успешной подготовки к семинару рекомендуется заранее ознакомиться с темой занятия и основными материалами, чтобы иметь возможность активно участвовать в обсуждении. Также полезно подготовить вопросы и идеи для обсуждения, что поможет глубже понять материал и продемонстрировать заинтересованность.

Консультации – структурированные встречи, на которых преподаватели предоставляют индивидуальную или групповую помощь в освоении учебного материала, обсуждении вопросов и решении проблем, возникающих в процессе обучения.

Консультации могут включать разъяснение сложных тем, подготовку к экзаменам и помощь в выполнении проектных работ, что способствует более глубокому пониманию предмета и улучшению академической успеваемости.

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.

Мини-проект — это краткосрочная практическая работа в малых группах на семинарах, направленная на применение теоретических знаний для решения конкретных задач.

В ходе мини-проекта студенты проводят анализ проблемы, разрабатывают техническое решение и представляют результаты в виде отчета, модели или презентации. Такой формат способствует развитию навыков самостоятельной работы, критического мышления и командного взаимодействия.

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) «Хранилища больших данных»

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	

Десятибалльная оценка	Пятибалльная оценка	Общая характеристика результата обучения по дисциплине (модулю)
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, касающиеся сути вопросов. Студент способен решать только самые простые задачи и владеет лишь минимальным набором методов исследования.
4	Удовлетворительно	
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) «Хранилища больших данных» оценивается следующим образом:

Активность	Вес	Количество	Описание
Домашние задания	40%	5	Набор задач по темам недели
Экзамен	60%	1	Мини-проект — краткосрочная практическая работа в малых группах

Формула расчёта итоговой оценки по дисциплине (модулю) «Хранилища больших данных»: «0,4 × среднее за домашние задания + 0,6 × экзамен».

Текущий контроль успеваемости обучающихся по дисциплине (модулю)

Примерные домашние задания

Домашнее задание по теме «Введение в DWH».

Напишите короткий аналитический обзор (1–2 страницы), в котором сравните архитектуру хранилища данных (DWH) и data lake на примере бизнеса электронной коммерции: объясните, зачем компании нужно DWH, опишите назначение слоёв staging, core и marts, приведите по одному примеру типового сценария использования и анти-паттерна (например, "spaghetti pipelines" или отсутствие слоя core).

Домашнее задание по теме «Моделирование».

На основе бизнес-вопроса *«Какова выручка по категориям товаров в разрезе регионов и времени?»* спроектируйте модель данных в формате "звезда": определите таблицы фактов (указав гранулярность — например, заказ-позиция) и измерений (товар, категория, регион, время), обоснуйте выбор гранулярности и нарисуйте схему (можно в виде ASCII-диаграммы или в текстовом описании связей). Кратко объясните, чем ваша модель отличается от схемы "снежинка" и почему вы выбрали именно "звезду".

Домашнее задание по теме «Качество и governance».

Для одной из таблиц в слое core (fct_orders) составьте:

— 3 проверки качества данных (например, not null, uniqueness, referential integrity);

— схематичное описание lineage: от источника до core (какие слои, какие трансформации);

— фрагмент документации (в формате Markdown): описание таблицы, полей, владельца данных (data owner), срока хранения. Объясните, как наличие таких практик помогает предотвратить ошибки в аналитике.

Домашнее задание по теме «Доступы и безопасность».

Опишите модель управления доступом в DWH для трёх ролей: аналитик, data engineer, финансовый директор. Укажите, какие таблицы и поля им доступны, где необходимо применить маскирование (например, для email или ЗП), и какие действия должны логироваться (аудит). Дополнительно: предложите политику хранения данных для слоя staging (например, 30 дней) и core (бессрочно), обоснуйте свой выбор с точки зрения безопасности и compliance.

Примеры заданий и критерии оценивания мини-проекта.

Задачи и этапы выполнения проекта:

Разработать концептуальный проект хранилища данных, включающий:

1. Описание архитектуры слоёв (staging, core, marts) и обоснование их использования.
2. Модель данных в слое core по схеме "звезда": определите таблицы фактов (с указанием гранулярности) и измерений.
3. Реализацию SCD Type 2 для одного из измерений (например, dim_users или dim_courses) и описание стратегии инкрементальной загрузки.
4. Проектирование витрины данных (data mart) для аналитики продаж — с агрегацией по дням и категориям курсов.
5. Перечень проверок качества данных, краткая схема lineage от источника до витрины и фрагмент документации.
6. Описание политики доступа: роли (аналитик, администратор, маркетолог), маскирование персональных данных, политика хранения для staging (30 дней).
7. Презентация решения (5–7 слайдов) или структурированный отчёт (в формате Markdown/PDF).

Критерии оценивания.

1. Соответствие архитектуры DWH современным практикам — корректное

использование слоёв (staging/core/marts), обоснование их назначения и избегание типовых анти-паттернов.

2. Качество модели данных — правильный выбор гранулярности, выделение фактов и измерений, обоснованное применение схемы "звезда", логичная структура таблиц.
3. Учёт истории изменений и механизмов загрузки — корректная реализация SCD Type 2, описание обработки инкрементов, поздних данных или backfill.
4. Практичность и эффективность витрины данных — релевантность агрегатов бизнес-задаче, учёт производительности запросов, логичное выделение marts.
5. Обеспечение качества данных и управление данными (governance) — наличие проверок качества, описание lineage, наличие документации и указание владельцев данных.
6. Организация доступа и безопасность — адекватное распределение ролей, применение маскирования, наличие аудита и обоснованные политики хранения.
7. Командная работа и презентация результата — ясность изложения, структурированность отчёта или презентации, логичная аргументация решений, вклад всех участников группы.

Задания для промежуточной аттестации по дисциплине (модулю)

№ п/п	Задание	Ответ	Компетенция
1	В хранилище данных выделяют три основных слоя. Назовите слой, в котором хранятся необработанные данные, поступившие из источников, и укажите их количество:	1 staging	УК-1
2	При моделировании данных в DWH используются таблицы фактов и измерений. Укажите, сколько таблиц фактов обычно присутствует в схеме "звезда", и назовите тип таблицы, содержащей бизнес-метрики:	1 фактов	УК-1
3	Для отслеживания изменений в измерениях применяют SCD. Укажите тип SCD, который позволяет хранить историю изменений с помощью дат начала и окончания действия записи, и назовите его номер:	2 Type	УК-1
4	Вы проектируете DWH для анализа продаж. Определите, сколько уровней агрегации целесообразно выделить в витрине данных для ежедневного отчёта по выручке, и назовите слой, отвечающий за эти витрины:	1 marts	УК-2
5	Вам нужно организовать загрузку данных по пользователям, которые могут менять регион. Укажите, сколько методов инкрементальной загрузки вы будете использовать одновременно при наличии поздних данных, и назовите общий принцип обновления таких данных:	1 backfill	УК-2
6	При нарушении качества данных вы внедряете проверки. Укажите, сколько типов базовых проверок (например, на отсутствие NULL) вы реализуете в слое core, и назовите процесс отслеживания происхождения данных:	1 lineage	УК-2
7	В модели "звезда" таблица фактов связана с измерениями. Укажите количество внешних ключей в строке таблицы фактов для четырёх измерений и назовите тип такой связи:	4 foreign	ОПК-1
8	При расчёте выручки по заказам вы используете агрегатную функцию. Укажите её порядковый номер среди стандартных SQL-агрегатов (например, COUNT=1, SUM=2) и назовите саму функцию:	2 SUM	ОПК-1

9	При работе с временными данными вы используете тип, позволяющий хранить дату и время. Укажите номер этого типа данных в списке: DATE=1, TIME=2, DATETIME=3, и назовите его:	3 DATETIME	ОПК-1
10	Вы выбираете технологию для построения DWH. Укажите, сколько из перечисленных — Apache Spark, Tableau, Power BI — относятся к инструментам обработки данных, и назовите одну из них:	1 Spark	ОПК-5
11	Для визуализации данных из витрин используются BI-инструменты. Укажите, сколько таких инструментов из списка: SQL, dbt, Looker — подходят для визуализации, и назовите один из них:	1 Looker	ОПК-5
12	При построении пайплайнов вы используете оркестратор. Укажите, сколько из перечисленных — Airflow, Kafka, Docker — являются системами оркестрации, и назовите одну из них:	1 Airflow	ОПК-5
13	Вы разрабатываете витрину данных для аналитиков. Укажите, сколько уровней доступа (например, чтение, запись) будет у роли "аналитик", и назовите тип доступа, который она получит:	1 чтение	ПК-2
14	Для защиты персональных данных вы применяете технику сокрытия значений. Укажите, сколько способов маскирования (например, хеширование, замена звёздочками) вы внедрите в слое marts, и назовите саму технику:	1 маскирование	ПК-2
15	При работе с данными вы соблюдаете политики хранения. Укажите, сколько дней хранятся временные данные в слое staging при стандартной политике, и назовите этот слой:	30 staging	ПК-2
16	Вы разрабатываете ETL-процесс загрузки данных. Укажите количество этапов в классической схеме ETL и назовите первый этап:	3 extract	ПК-4
17	Вы пишете SQL-скрипт для создания таблицы измерений. Укажите, сколько основных типов данных (например, INTEGER, VARCHAR, DATE) вы используете в таблице с 5 атрибутами, и назовите один из них:	3 VARCHAR	ПК-4
18	Вы документируете модель данных. Укажите, сколько обязательных элементов (название таблицы, описание полей, владелец данных) включает контракт данных, и назовите один из них:	3 документация	ПК-4