

Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Андрей Драгомирович Хлутков
Должность: директор
Дата подписания: 05.08.2026 12:36:35
Уникальный программный ключ:
880f7c07c583b07b775f6604b639281b15a09f12

Федеральное государственное бюджетное образовательное учреждение
высшего образования

**РОССИЙСКАЯ АКАДЕМИЯ НАРОДНОГО ХОЗЯЙСТВА и
ГОСУДАРСТВЕННОЙ СЛУЖБЫ при ПРЕЗИДЕНТЕ РОССИЙСКОЙ
ФЕДЕРАЦИИ**

СЕВЕРО-ЗАПАДНЫЙ ИНСТИТУТ УПРАВЛЕНИЯ

Факультет среднего профессионального образования

УТВЕРЖДЕНА
решением цикловой (методической)
комиссии общепрофессиональных
дисциплин и профессиональных
модулей специальностей 09.02.07
Информатика и вычислительная
техника
Протокол от 31.10.2025 № 2

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

МДК.02.01. Инжиниринг данных

Специальность – 09.02.12 Техническая эксплуатация и сопровождение информационных систем

Профиль – на базе основного общего образования

Квалификация – специалист по технической эксплуатации и сопровождению информационных систем

Форма обучения – очная

Год набора – 2026

Санкт-Петербург 2025 год

Автор-составитель: Бурылов Василий Сергеевич, преподаватель ФСПО СЗИУ РАНХиГС

СОДЕРЖАНИЕ

1. Общие положения	4
1.1. Область применения программы	4
1.2. Место дисциплины в структуре основной профессиональной образовательной программы	4
1.3. Цели и задачи междисциплинарного курса.....	4
1.4. Планируемые результаты обучения по дисциплине	4
2. Структура и содержание дисциплины	12
2.1. Объем учебной дисциплины и виды работ	12
2.2. Тематический план и содержание дисциплины	12
2.3. Регламент распределения видов работ по дисциплине с ДОТ	14
3. Материалы текущего контроля успеваемости и промежуточной аттестации обучающихся	15
3.1. Формы и методы текущего контроля успеваемости обучающихся и промежуточной аттестации.....	15
3.2. Оценочные средства текущего контроля успеваемости обучающихся	16
3.3. Оценочные средства промежуточной аттестации обучающихся	22
4. Методические указания для обучающихся по освоению дисциплины	24
5. Учебная литература и ресурсы информационно-телекоммуникационной сети «Интернет»	28
6. Материально-техническая база, информационные технологии, программное обеспечение и информационные справочные системы	30

1 Общие положения

1.1 Область применения программы

Рабочая программа междисциплинарного курса «Инжиниринг данных» является частью профессиональной подготовки обучающихся направления подготовки 09.02.12 «Техническая эксплуатация и сопровождение информационных систем».

1.2 Место дисциплины в структуре основной профессиональной образовательной программы

Междисциплинарный учебный курс «Инжиниринг данных» является частью профессионального модуля «Конфигурирование аналитических решений», который входит в профессиональный цикл подготовки.

Для успешного освоения междисциплинарного курса предполагается освоение таких предшествующих дисциплин и модулей, как «Математический аппарат в отрасли информационных технологий», «Операционные системы и среды», «Основы алгоритмизации и программирования», «Основы информационной безопасности», «Основы работы с информацией», «Математический анализ», «Введение в науку о данных. SQL и Python». Полученные в результате освоения дисциплины знания, умения и практические навыки необходимы и будут полезны для последующего изучения дисциплины модуля «Конфигурирование аналитических решений».

Дисциплина изучается на 3 курсе в 5 семестре.

1.3 Цели и задачи междисциплинарного курса

Цель междисциплинарного курса «Инжиниринг данных» – изучение и практическая реализация сбора, хранения и очистки данных для последующего анализа.

1.4 Планируемые результаты обучения по дисциплине

Перечень общих компетенций

Код и наименование компетенции	Умения	Знания
ОК 01 Выбирать способы решения задач профессиональной деятельности применительно к различным контекстам	- распознавать задачу и/или проблему в профессиональном и/или социальном контексте, анализировать и выделять её составные части; - определять этапы решения задачи, составлять план действия, реализовывать составленный план, определять необходимые ресурсы; - выявлять и эффективно искать информацию, необходимую для решения задачи и/или проблемы; - владеть актуальными методами работы в профессиональной и смежных сферах;	- актуальный профессиональный и социальный контекст, в котором приходится работать и жить; - структура плана для решения задач, алгоритмы выполнения работ в профессиональной и смежных областях; - основные источники информации и ресурсы для решения задач и/или проблем в профессиональном и/или социальном контексте; - методы работы в профессиональной и смежных сферах;

Код и наименование компетенции	Умения	Знания
	- оценивать результат и последствия своих действий (самостоятельно или с помощью наставника);	- порядок оценки результатов решения задач профессиональной деятельности;
ОК 02 Использовать современные средства поиска, анализа и интерпретации информации, и информационные технологии для выполнения задач профессиональной деятельности	- определять задачи для поиска информации, планировать процесс поиска, выбирать необходимые источники информации; - выделять наиболее значимое в перечне информации, структурировать получаемую информацию, оформлять результаты поиска; - оценивать практическую значимость результатов поиска; - применять средства информационных технологий для решения профессиональных задач; - использовать современное программное обеспечение в профессиональной деятельности; - использовать различные цифровые средства для решения профессиональных задач;	- номенклатура информационных источников, применяемых в профессиональной деятельности; - приемы структурирования информации; - формат оформления результатов поиска информации; - современные средства и устройства информатизации, порядок их применения и программное обеспечение в профессиональной деятельности, в том числе цифровые средства;
ОК 03 Планировать и реализовывать собственное профессиональное и личностное развитие, предпринимательскую деятельность в профессиональной сфере, использовать знания по правовой и финансовой грамотности в различных жизненных ситуациях	- определять актуальность нормативно-правовой документации в профессиональной деятельности; - применять современную научную профессиональную терминологию; - определять и выстраивать траектории профессионального развития и самообразования; - выявлять достоинства и недостатки коммерческой идеи; - определять инвестиционную привлекательность коммерческих идей в рамках профессиональной деятельности, выявлять источники финансирования; - презентовать идеи открытия собственного дела в профессиональной деятельности; - определять источники достоверной правовой информации; - составлять различные правовые документы; - находить интересные проектные идеи, грамотно их формулировать и документировать; - оценивать жизнеспособность проектной идеи, составлять план проекта;	- содержание актуальной нормативно-правовой документации; - современная научная и профессиональная терминология; - возможные траектории профессионального развития и самообразования; - основы предпринимательской деятельности, правовой и финансовой грамотности; - правила разработки презентации; - основные этапы разработки и реализации проекта;
ОК 04	- организовывать работу коллектива и команды;	- психологические основы деятельности коллектива;

Код и наименование компетенции	Умения	Знания
Эффективно взаимодействовать и работать в коллективе и команде	- взаимодействовать с коллегами, руководством, клиентами в ходе профессиональной деятельности;	- психологические особенности личности;
ОК 05 Осуществлять устную и письменную коммуникацию на государственном языке Российской Федерации с учетом особенностей социального и культурного контекста	- грамотно излагать свои мысли и оформлять документы по профессиональной тематике на государственном языке; - проявлять толерантность в рабочем коллективе;	- правила оформления документов; - правила построения устных сообщений; - особенности социального и культурного контекста;
ОК 06 Проявлять гражданско- патриотическую позицию, демонстрировать осознанное поведение на основе традиционных общечеловеческих ценностей, в том числе с учетом гармонизации межнациональных и межрелигиозных отношений, применять стандарты антикоррупционного поведения	- проявлять гражданско- патриотическую позицию; - демонстрировать осознанное поведение; - описывать значимость своей специальности; - применять стандарты антикоррупционного поведения;	- сущность гражданско- патриотической позиции; - традиционных общечеловеческих ценностей, в том числе с учетом гармонизации межнациональных и межрелигиозных отношений; - значимость профессиональной деятельности по специальности; - стандарты антикоррупционного поведения и последствия его нарушения;
ОК 07 Содействовать сохранению окружающей среды, ресурсосбережению, применять знания об изменении климата, принципы бережливого производства, эффективно действовать в чрезвычайных ситуациях	- соблюдать нормы экологической безопасности; - определять направления ресурсосбережения в рамках профессиональной деятельности по специальности; - организовывать профессиональную деятельность с соблюдением принципов бережливого производства; - организовывать профессиональную деятельность с учетом знаний об изменении климатических условий региона; - эффективно действовать в чрезвычайных ситуациях;	- правила экологической безопасности при ведении профессиональной деятельности; - основные ресурсы, задействованные в профессиональной деятельности; - пути обеспечения ресурсосбережения; - принципы бережливого производства; - основные направления изменения климатических условий региона; - правила поведения в чрезвычайных ситуациях;
ОК 08 Использовать средства физической культуры для сохранения и укрепления здоровья в процессе профессиональной деятельности и	- использовать физкультурно- оздоровительную деятельность для укрепления здоровья, достижения жизненных и профессиональных целей; - применять рациональные приемы двигательных функций в профессиональной деятельности;	- роль физической культуры в общекультурном, профессиональном и социальном развитии человека; - основы здорового образа жизни; - условия профессиональной деятельности и зоны риска

Код и наименование компетенции	Умения	Знания
поддержания необходимого уровня	- пользоваться средствами профилактики перенапряжения, характерными для данной специальности;	физического здоровья для специальности; - средства профилактики перенапряжения;
ОК 09 Пользоваться профессиональной документацией на государственном и иностранном языках	- понимать общий смысл четко произнесенных высказываний на известные темы (профессиональные и бытовые), понимать тексты на базовые профессиональные темы; - участвовать в диалогах на знакомые общие и профессиональные темы; - строить простые высказывания о себе и о своей профессиональной деятельности; - кратко обосновывать и объяснять свои действия (текущие и планируемые); - писать простые связные сообщения на знакомые или интересующие профессиональные темы;	- правила построения простых и сложных предложений на профессиональные темы; - основные общепотребительные глаголы (бытовая и профессиональная лексика); - лексический минимум, относящийся к описанию предметов, средств и процессов профессиональной деятельности; - особенности произношения, правила чтения текстов профессиональной направленности;

Перечень профессиональных компетенций

Код и наименование компетенции	Умения	Знания	Владеть навыками
ПК 2.1 Выполнять подготовку данных для проведения аналитических работ	– Определять требования к поставщикам данных из гетерогенных источников – Осуществлять взаимодействие с внутренними и внешними поставщиками данных из гетерогенных источников – Разрабатывать и оценивать модели больших данных – Использовать инструментальные средства для извлечения, преобразования, хранения и обработки данных из разнородных источников, в том	– Возможности имеющейся у исполнителя методологической и технологической инфраструктуры анализа больших данных – Предметную область анализа – Теоретические и прикладные основы анализа больших данных – Современные методы и инструментальные средства анализа больших данных – Современный опыт использования анализа больших данных – Типы больших данных: метаданные, полуструктурированные, структурированные, неструктурированные – Виды источников данных: созданные	– Определения источников больших данных для анализа, идентификация внешних и внутренних источников данных для проведения аналитических работ – Получения и фильтрации больших объемов данных из гетерогенных источников – Извлечения, проверки и очистки больших объемов данных из гетерогенных источников – Агрегации и разработки представления больших объемов данных из гетерогенных источников – Оценки соответствия набора данных

Код и наименование компетенции	Умения	Знания	Владеть навыками
	числе в режиме реального времени – Производить очистку данных для проведения аналитических работ – Проводить интеграцию и преобразование больших объемов данных – Оценивать соответствие наборов данных задачам анализа больших данных – Оценивать стоимость данных для проведения аналитических работ	человеком, созданные машинами – Источники информации, в том числе информации, необходимой для обеспечения деятельности в предметной области заказчика исследования – Методы извлечения информации и знаний из гетерогенных, мультиструктурированных, неструктурированных источников, в том числе при потоковой обработке – Российские и международные стандарты информационной безопасности – Современную технологическую инфраструктуру высокопроизводительных и распределенных вычислений – Режимы получения и обработки данных, поддержка режима реального времени – Технологии хранения и обработки больших данных в организации: базы данных, хранилища данных, распределенная и параллельная обработка данных, вычисления в оперативной памяти – Облачные технологии, облачные сервисы – Методы оценки временных и стоимостных характеристик технологий больших данных – Технологии межличностной и групповой коммуникации в деловом взаимодействии, основы конфликтологии – Правила деловой переписки	предметной области и задачам аналитических работ

Код и наименование компетенции	Умения	Знания	Владеть навыками
ПК 2.2 Строить статистические и математические модели для систем анализа данных	– Формулировать математические модели для реальных задач – Обрабатывать пропущенные данные и аномалии. – Представлять данные и результаты анализа с помощью графиков и диаграмм. – Самостоятельно изучать новые инструменты и технологии.	– Основы теории вероятностей и ее применения в статистике. – Матрицы, векторы и операции с ними. – Основы математического анализа для понимания непрерывных функций. – Методы анализа экономических данных. – Специфику и особенности предметной области анализа данных (финансы, медицина, маркетинг и т.д.) – Методы оптимизации – Основы управления проектами для успешного выполнения задач в срок.	– Понимания основных статистических понятий. – Умения проводить описательную и инференциальную статистику. – Опыта работы с языками программирования (Python, R, MATLAB). – Знания библиотек для анализа данных (NumPy, pandas, scikit-learn и т.д.). – Умения собирать, обрабатывать и визуализировать данные. – Опыта работы с инструментами визуализации (Matplotlib, Seaborn, Tableau). – Понимания основных алгоритмов машинного обучения. – Умения работать в команде и эффективно взаимодействовать с другими участниками проекта. – Способности анализировать и интерпретировать результаты.
ПК 2.3 Конфигурировать информационные системы анализа данных	– Видеть взаимосвязи между различными компонентами информационных систем и понимать, как они взаимодействуют. – Проводить анализ данных и интерпретировать результаты для принятия обоснованных решений. – Выявлять и устранять проблемы	– Архитектуру информационных систем, включая клиент-серверные модели, облачные технологии и распределенные системы. – Статистику и методы анализа данных, необходимых для конфигурации и использования систем. – Принципы обеспечения безопасности данных и соблюдения нормативных требований. – Процессы извлечения, преобразования и загрузки	– Работы с SQL и NoSQL базами данных, включая создание запросов, управление данными и оптимизацию производительности. – Создания визуализаций с помощью инструментов, таких как Tableau, Power BI или Matplotlib. – Тестирования систем и отладки конфигураций для обеспечения их корректной работы

Код и наименование компетенции	Умения	Знания	Владеть навыками
	в конфигурации систем и их интеграции. – Четко и эффективно общаться с командой и заинтересованными сторонами, объясняя технические аспекты и результаты анализа. – Планировать и управлять проектами по внедрению информационных систем. – Писать код на языках программирования, таких как Python, R, или Java, для автоматизации процессов и настройки систем. – Устанавливать и настраивать программное обеспечение для анализа данных и BI (Business Intelligence) систем	данных в системы для анализа. – Технологии и инструменты, такие как Hadoop, Spark или Apache Kafka, для работы с большими данными.	
ПК 2.4 Формировать визуальные решения на основе информационных систем анализа данных	– Оценивать данные и визуализации, чтобы выбрать наиболее подходящие методы и подходы для представления информации. – Анализировать данные и выявлять ключевые метрики и тренды, которые должны быть визуализированы. – Четко и понятно представлять результаты визуализаций как техническим, так и	– Статистические методы и принципы анализа данных, необходимые для интерпретации результатов. – Основные принципы визуализации, такие как выбор правильных типов графиков, использование цвета и композиции. – Бизнес-аналитику и инструменты, которые помогают в анализе данных и создании отчетов. – Этику, связанную с обработкой и визуализацией данных, включая	– Программирование на языках, таких как Python или R, для обработки данных и создания кастомизированных визуализаций.

Код и наименование компетенции	Умения	Знания	Владеть навыками
	нетехническим пользователям. – Выявлять и решать проблемы, связанные с данными и визуализациями, например, выбросы или недостающие данные. – Эффективно работать в команде с аналитиками, разработчиками и другими заинтересованными сторонами для создания комплексных визуальных решений. – Быстро адаптироваться к новым инструментам, технологиям и изменениям в требованиях проекта. – Создавать информативные и эстетически привлекательные визуализации с использованием инструментов, таких как Tableau, Power BI, QlikView, D3.js и Matplotlib. – Извлекать данные из SQL и NoSQL баз данных, а также работать с API для получения данных.	конфиденциальность и безопасность информации. – Современные тренды и лучшие практики в области визуализации данных и анализа. – Основы UX/UI дизайна для создания удобных и интуитивно понятных интерфейсов визуализации.	

2 Структура и содержание дисциплины

2.1 Объем учебной дисциплины и виды работ

Виды учебной работы	Объем учебной работы, час.
Учебная нагрузка обучающихся всего, в том числе:	70
лекции	12
практические занятия	36
курсовой проект	20
самостоятельная работа обучающихся	-
консультации	2
промежуточная аттестация	-
Форма промежуточной аттестации	Зачет с оценкой

2.2 Тематический план и содержание дисциплины

№ п/п	Наименование тем (разделов)	Содержание тем (разделов)	Распределение часов			Формируемые компетенции	Формы текущего контроля
			Л	ПР	СРС		
1.	Тема 1. Сбор, хранение и очистка данных	Содержание учебного материала Архитектура хранилищ данных: DWH, Data Lake, Data Mart. Типы источников: реляционные СУБД, API, файлы, стриминг Методы извлечения данных: SQL, Python, ETL-инструменты Проблемы качества данных: пропуски, дубликаты, аномалии Стандартизация и нормализация: схемы, форматы, кодировки Каталогизация и метаданные: Data Catalog, Data Lineage Практические занятия 1. Подключение к PostgreSQL, импорт из Excel, JSON и API 2. SQL-запросы для выборки и фильтрации 3. Очистка данных от пропусков, нормализация форматов дат 4. Преобразование данных в Pandas (Python) 5. Составление Data Dictionary и схемы зависимостей	4	12	-	ОК 01, ОК 02, ОК 03, ОК 04, ОК 05, ОК 06, ОК 07, ОК 08, ОК 09, ПК 2.1, ПК 2.2, ПК 2.3, ПК 2.4	О, ПЗ,Т
2.	Тема 2. ETL и трансформация данных	Содержание учебного материала Концепция ETL/ELT: Extract, Transform, Load Инструменты ETL: Apache NiFi, Airbyte, Talend, dbt (обзор) Написание трансформаций: SQL, Python (Pandas, PySpark)	4	12	-	ОК 01, ОК 02, ОК 03, ОК 04, ОК 05, ОК 06, ОК 07, ОК 08, ОК 09,	О, ПЗ,Т

№ п/п	Наименование тем (разделов)	Содержание тем (разделов)	Распределение часов			Форми- руемые компетен ции	Формы текущего контроля
			Л	ПР	СРС		
		Расчёт derived-полей, агрегация, оконные функции Планирование задач: Apache Airflow, cron Контроль качества и отладка пайплайнов Управление состоянием: StateFlow, LiveData, remember {} Локализация интерфейса и мультиязычность Практические занятия 1. Настройка простого ETL в Python: CSV → PostgreSQL 2. Использование оконных функций для расчёта метрик 3. Сценарий агрегации в dbt / Airflow DAG с расписанием 4. Отладка пайплайна с логированием и проверкой ошибок 5. Построение репозитория с пайплайнами и версионированием				ПК 2.1, ПК 2.2, ПК 2.3, ПК 2.4	
3.	Тема 3. Инженерия данных для аналитики и ML	Содержание учебного материала Подготовка данных под BI и ML: структурирование, типизация Обогащение: объединения, справочники, геоданные Особенности хранения временных рядов, событий и транзакций Фиче-инжиниринг: создание новых переменных, категориальных и числовых Хранение результатов моделей и логики предиктивной аналитики Документирование дата-пайплайнов и соблюдение SLA Практические занятия 1. Создание витрины данных для дашборда 2. Построение признаков на основе поведения пользователя 3. Интеграция модели прогноза в пайплайн (scikit-learn + Pandas) 4. Сбор и хранение событий в time-series формате 5. Визуализация lineage в dbt и документация по DataHub	4	14	-	ОК 01, ОК 02, ОК 03, ОК 04, ОК 05, ОК 06, ОК 07, ОК 08, ОК 09, ПК 2.1, ПК 2.2, ПК 2.3, ПК 2.4	О, ПЗ,Т

№ п/п	Наименование тем (разделов)	Содержание тем (разделов)	Распределение часов			Форми- руемые компетен- ции	Формы текущего контроля
			Л	ПР	СРС		
		Итого часов:	12	36			

2.3 Регламент распределения видов работ по дисциплине с ДОТ

Данная дисциплина реализуется с применением дистанционных образовательных технологий (ДОТ). Распределение видов учебной работы, форматов текущего контроля представлены в Таблице 2.3.

Таблица 2.3. — Распределение видов учебной работы и текущей аттестации

Вид учебной работы	Формат проведения
Лекционные занятия	Частично с применением ДОТ
Практические занятия	Частично с применением ДОТ
Курсовой проект	Частично с применением ДОТ
Текущий контроль	Частично с применением ДОТ
Промежуточная аттестация	Контактная аудиторная работа
Формы текущего контроля	Формат проведения
Тестирование	Частично с применением ДОТ
Доклады	Контактная аудиторная работа
Опрос	Контактная аудиторная работа
Практические задания	Частично с применением ДОТ

Доступ к системе дистанционных образовательных программ осуществляется каждым обучающимся самостоятельно с любого устройства на портале: <https://sziu-de.ranepa.ru> в соответствии с их индивидуальным паролем и логином к личному кабинету / профилю.

Текущий контроль, проводимый в системе дистанционного обучения, оцениваются как в системе дистанционного обучения, так и преподавателем вне системы.

Доступ к материалам лекций предоставляется в течение всего семестра по мере прохождения освоения программы. Доступ к каждому виду работ и количество попыток на выполнение задания предоставляется ограниченное время согласно регламенту дисциплины, опубликованному в системе дистанционного обучения. Преподаватель оценивает выполненные обучающимися работы не позднее 14 рабочих дней после окончания срока выполнения.

3 Материалы текущего контроля успеваемости и промежуточной аттестации обучающихся

3.1 Формы и методы текущего контроля успеваемости и промежуточной аттестации обучающихся

Формы текущего контроля успеваемости:

Опрос (О) позволяет выявить правильность ответа по содержанию, его последовательность, самостоятельность суждений и выводов, степень развития логического мышления.

Оценка	Критерии оценивания
«Отлично»	Ответ правильный по содержанию, логически выстроен и последователен. Студент демонстрирует самостоятельность суждений и выводов, свободно оперирует терминами, раскрывает суть понятий и их взаимосвязи. Проявляется высокий уровень развития логического мышления: студент способен анализировать, сопоставлять, приводить примеры и аргументировать позицию.
«Хорошо»	Ответ в целом правильный и достаточно последовательный, отражает понимание основных положений темы. Студент владеет терминологией, но может испытывать небольшие затруднения при раскрытии сложных взаимосвязей или аргументации. Допускаются незначительные неточности, которые студент способен исправить самостоятельно после наводящих вопросов.
«Удовлетворительно»	Ответ содержит основные сведения по теме, но отличается недостаточной последовательностью, фрагментарностью или слабой аргументацией. Студент знает базовые понятия, но испытывает трудности в раскрытии их взаимосвязей и применении на практике. Для устранения пробелов требуется руководство преподавателя.
«Неудовлетворительно»	Ответ неправильный или крайне неполный, отсутствует логическая структура, наблюдается путаница в терминах и понятиях. Студент не способен самостоятельно сформулировать выводы, не демонстрирует понимания сути

	темы. Требуется повторное изучение основных разделов дисциплины под руководством преподавателя.
--	---

Тестирование (Т) – задания, с вариантами ответов.

Оценка	Критерии оценивания
«Отлично»	Студент правильно ответил на 90–100 % вопросов теста. Ответы демонстрируют уверенное владение материалом, отсутствие ошибок в базовых и усложнённых заданиях.
«Хорошо»	Студент правильно ответил на 75–89 % вопросов. Допускаются отдельные неточности, не искажающие суть понятий; в целом материал усвоен, но есть пробелы в отдельных темах.
«Удовлетворительно»	Студент правильно ответил на 50–74 % вопросов. Усвоены базовые понятия, но имеются существенные пробелы; допускаются ошибки в применении правил и интерпретации условий заданий.
«Неудовлетворительно»	Студент правильно ответил менее чем на 50 % вопросов либо не представил тест на проверку. Проявлены значительные пробелы в знаниях, непонимание ключевых тем дисциплины.

Практическое задание (ПЗ) используется для закрепления теоретических знаний и отработки навыков и умений, способности применять знания при решении конкретных задач.

Оценка	Критерии оценивания
«Отлично»	Студент демонстрирует глубокое знание материала и свободно выполняет задание. Понимает взаимосвязь основных понятий темы, обосновывает выбранные методы решения, корректно интерпретирует результаты. Работа выполнена полностью, без ошибок, с соблюдением всех требований к оформлению и срокам сдачи.
«Хорошо»	Студент полностью знает материал и успешно выполняет предусмотренные задания. Допускает незначительные ошибки (неточность фактов, небольшие погрешности в расчётах или оформлении, стилистические неточности), которые не влияют

	на общий результат и могут быть быстро исправлены самостоятельно.
«Удовлетворительно»	Студент владеет основным материалом в объёме, необходимом для дальнейшего изучения дисциплины, и справляется с выполнением задания. Допускает погрешности в решении или оформлении, но обладает необходимыми знаниями для их устранения под руководством преподавателя. Работа в целом соответствует требованиям, но нуждается в доработке.
«Неудовлетворительно»	Студент имеет существенные пробелы в знании основного материала, не справляется с выполнением задания или допускает серьёзные ошибки, искажающие результат. Нуждается в повторении основных разделов курса под руководством преподавателя; работа не соответствует требованиям либо сдана с грубыми нарушениями сроков и формата.

Формы текущего контроля

№	Название темы	Формы текущего контроля успеваемости
1.	Тема 1. Сбор, хранение и очистка данных	Т, ПЗ, О
2.	Тема 2. ETL и трансформация данных	Т, ПЗ, О
3.	Тема 3. Инженерия данных для аналитики и ML	Т, ПЗ, О

Примечание. В столбце «Форма текущего контроля успеваемости, промежуточной аттестации» перечисляются все используемые в учебном процессе по данной дисциплине формы контроля освоения материала. (Т – тестирование; ПЗ – практическое задание, О – опрос).

Промежуточная аттестация проводится в виде зачета с оценкой. Условием допуска обучающегося к промежуточной аттестации является освоение материалов учебной дисциплины в объеме не менее 75 %, определенное по результатам систематического текущего контроля.

Критерии оценивания зачёта с оценкой:

Оценка «отлично» ставится, если:

- студент полно излагает материал (отвечает на вопрос), дает правильное определение основных понятий;
- обнаруживает понимание материала, может обосновать свои суждения, применить знания на практике, привести необходимые примеры не только из учебника, но и самостоятельно составленные;
- излагает материал последовательно и правильно с точки зрения норм литературного языка.

Оценка «хорошо» ставится, если студент дает ответ, удовлетворяющий тем же требованиям, что и для оценки «отлично», но допускает 1–2 ошибки, которые сам же исправляет, и 1–2 недочета в последовательности и языковом оформлении излагаемого.

Оценка «удовлетворительно» ставится, если:

- студент обнаруживает знание и понимание основных положений данной темы, но излагает материал неполно и допускает неточности в определении понятий или формулировке правил;
- не умеет достаточно глубоко и доказательно обосновать свои суждения и привести свои примеры;
- излагает материал непоследовательно и допускает ошибки в языковом оформлении излагаемого.

Оценка «неудовлетворительно» ставится, если:

- студент обнаруживает незнание большей части соответствующего вопроса, допускает ошибки в формулировке определений и правил, искажающие их смысл, беспорядочно и неуверенно излагает материал.

3.2 Материалы текущего контроля успеваемости обучающихся

Примеры типовых заданий для практических работ

Практическая работа 1.1. Подключение к PostgreSQL, импорт из Excel, JSON и API

Задание:

1. Подключение к БД
 - Создать БД import_db
 - Создать таблицу employees
 - Написать скрипт подключения на Python
2. Импорт из Excel
 - Загрузить данные из employees.xlsx

- Преобразовать типы данных
- Вставить записи в таблицу
- 3. Импорт из JSON
 - Прочитать файл skills.json
 - Обновить записи сотрудников
 - Добавить навыки в JSON-поле
- 4. Импорт через API
 - Получить данные с публичного API
 - Преобразовать структуру
 - Загрузить в БД
- 5. Итоговая проверка
 - Вывести статистику импорта
 - Проверить целостность данных
 - Экспорт результата в CSV

Практические работы 2.2 Использование оконных функций для расчёта метрик

Задание:

Напишите SQL-запросы с оконными функциями для решения задач:

Ранжирование

Для каждого пользователя выведите 3 самых дорогих заказа (используйте ROW_NUMBER())

Скользящее среднее

Рассчитайте скользящее среднее выручки за 7 дней для каждого продукта

Доля в категории

Для каждого заказа посчитайте, какой процент от общей суммы категории он составляет

Сравнение с предыдущим

Выведите разницу в сумме заказа с предыдущим заказом того же пользователя

Накопительный итог

Рассчитайте накопительную сумму выручки по месяцам

Сравнение со средним

Отметьте заказы, сумма которых выше средней по категории

Групповое ранжирование

Определите лучший (по сумме) заказ в каждой категории за каждый месяц

Практические работы 3.2 Использование оконных функций для расчёта метрик

Задание:

Создать набор признаков для прогнозирования оттока клиентов на основе их цифрового следа.

Исходные данные (таблица user_sessions):

user_id, session_start, session_duration, pages_viewed, actions_count, device_type

Задачи:

Временные признаки:

Активность по часам дня (ночной/дневной пользователь)

Количество дней с последней сессии

Частота сессий за последние 7/30 дней

Поведенческие признаки:

Средняя длительность сессии

Тренд активности (растёт/падает за неделю)

Любимый тип устройства (mobile/desktop)

Сложные признаки:

Стабильность расписания (стандартное отклонение времени сессий)

Интенсивность кликов (actions_count / session_duration)

Сессионная целостность (отношение коротких (<1 мин) сессий ко всем)

Агрегации:

Суммарное время в приложении за месяц

Максимальное количество просмотренных страниц за сессию

Требования к результату:

SQL-запрос или Python-скрипт с генерацией признаков

Таблица user_features с 10+ признаками

Описание 3 наиболее информативных признаков

Примеры тестовых заданий

Вопросы к теме 1.1. Какой метод сбора данных характеризуется минимальным вмешательством исследователя и регистрацией событий в их естественном протекании?

А) Эксперимент

Б) Наблюдение

В) Опрос

Г) Фокус-группа

2. Что такое «грязные данные» (dirty data)?

А) Данные, полученные неэтичным путём

Б) Данные, хранящиеся на устаревших носителях

В) Данные с ошибками, противоречиями и несоответствиями

Г) Зашифрованные данные

3. Какой принцип хранения данных подразумевает, что данные должны быть доступны авторизованным пользователям в любое время?

А) Конфиденциальность (Confidentiality)

Б) Целостность (Integrity)

В) Доступность (Availability)

Г) Масштабируемость (Scalability)

Вопросы к теме 2.

1. Как расшифровывается аббревиатура ETL?

А) Extract, Translate, Load

Б) Extract, Transform, Load

В) Evaluate, Transform, Link

Г) Extract, Transfer, Load

2. На каком этапе ETL-процесса обычно выполняются такие операции, как очистка от дубликатов, приведение к единому формату и обогащение данных?

А) Extract (Извлечение)

Б) Transform (Трансформация)

В) Load (Загрузка)

Г) На всех этапах равномерно

3. Какая из перечисленных операций НЕ является типичной операцией трансформации данных?

А) Агрегация (например, SUM, AVG)

Б) Соединение таблиц (JOIN)

В) Извлечение данных из API

Г) Фильтрация строк по условию

Вопросы к теме 3.

1. Какой тип системы хранения данных обычно служит единым источником истины для сырых, необработанных данных в их исходном формате?

А) Data Warehouse (Хранилище данных)

Б) Data Lake (Озеро данных)

В) Операционная база данных (OLTP)

Г) Feature Store (Хранилище признаков)

2. Что такое «пайплайн данных» (Data Pipeline)?

А) Графический интерфейс для построения отчетов

Б) Автоматизированная последовательность шагов по перемещению и преобразованию данных

В) Метод сжатия данных для экономии места

Г) Система управления доступом к данным

3. Для чего в ML используется «Feature Engineering» (конструирование признаков)?

А) Для выбора наиболее красивых графиков для дашборда

Б) Для создания из сырых данных новых переменных (признаков), которые лучше описывают закономерности для модели

В) Для индексации таблиц в базе данных

Г) Для автоматического тюнинга гиперпараметров модели

3.3 Оценочные средства промежуточной аттестации обучающихся

Примерная тематика курсовых проектов (курсовых работ)

1. Построение аналитической системы для интернет-магазина
2. Разработка витрины данных и дашборда для HR-аналитики
3. Анализ пользовательской активности и прогноза churn
4. Построение BI-платформы для службы логистики
5. Разработка системы анализа эффективности рекламы
6. Построение RFM-модели и визуализация сегментов
7. Интеграция BI и Python: аналитика для финтех-сервиса
8. Анализ клиентской базы и построение рекомендаций
9. Конфигурирование ETL и построение витрины в dbt
10. Анализ финансовых показателей компании и визуализация
11. BI-дешборд для руководства на основе реальных KPI
12. Анализ продаж и построение воронки в Power BI
13. Прогнозирование спроса по регионам и товарам
14. Аналитика загруженности склада с автоматическим обновлением

15. Построение системы мониторинга отклонений
16. Интеграция API внешних данных в BI-систему
17. Анализ текучести персонала и визуализация в Tableau
18. Разработка предиктивной аналитики и её визуализация
19. Построение панели контроля бизнес-метрик для ритейла
20. BI-решение для мониторинга отзывов клиентов

Вопросы для подготовки к зачету с оценкой

1. Архитектура хранилищ данных: DWH
2. Архитектура хранилищ данных: Data Lake
3. Архитектура хранилищ данных: Data Mart.
4. Типы источников: реляционные СУБД
5. Типы источников: API
6. Типы источников: файлы
7. Типы источников: стриминг
8. Методы извлечения данных: SQL
9. Методы извлечения данных: Python
10. Методы извлечения данных: ETL-инструменты
11. Проблемы качества данных: пропуски, дубликаты, аномалии
12. Стандартизация и нормализация: схемы, форматы, кодировки
13. Каталогизация и метаданные: Data Catalog
14. Каталогизация и метаданные: Data Lineage
15. Концепция ETL/ELT: Extract, Transform, Load
16. Инструменты ETL: Apache NiFi
17. Инструменты ETL: Airbyte
18. Инструменты ETL: Talend, dbt (обзор)
19. Написание трансформаций: SQL

20. Написание трансформаций: Python (Pandas, PySpark)
21. Планирование задач: Apache Airflow, cron
22. Контроль качества и отладка пайплайнов
23. Управление состоянием: StateFlow, LiveData, remember {}
24. Локализация интерфейса и мультиязычность
25. Подготовка данных под BI и ML: структурирование, типизация
26. Обогащение: объединения, справочники, геоданные
27. Особенности хранения временных рядов, событий и транзакций
28. Фиче-инжиниринг: создание новых переменных, категориальных и числовых
29. Хранение результатов моделей и логики предиктивной аналитики
30. Документирование дата-пайплайнов и соблюдение SLA

4 Методические указания для обучающихся по освоению дисциплины

Приступая к изучению междисциплинарного курса «Инжиниринг данных», студент должен ознакомиться с содержанием данной Рабочей программы с тем, чтобы иметь четкое представление о своей работе. Изучение курса осуществляется на основе выданных студенту преподавателем рекомендаций по выполнению всех заданий, предусмотренных учебным планом и программой.

В первую очередь необходимо уяснить цель и задачи изучаемого курса, оценить объем материала, познакомиться с предложенной и подобрать основную и дополнительную литературу, выявить наиболее важные проблемы, стоящие по вопросам изучаемой дисциплины.

Выполнение заданий осуществляется в соответствии с учебным планом и программой. Они должны выполняться в соответствии с методическими рекомендациями, выданными преподавателем, и представлены в установленные преподавателем сроки.

Работая с учебниками и учебными пособиями, целесообразно законспектировать тот материал, который не сообщался студентам на лекциях.

На занятиях лекционного и практического характера студентам для работы требуется тетрадь для записи лекций и заданий.

Для успешного овладения программой дисциплины необходимо выполнять следующие требования:

- посещать все лекционные и практические занятия;

— все рассматриваемые на лекциях и практических занятиях темы и вопросы обязательно фиксировать в тетради;

— в случае пропуска занятий по каким-либо причинам необходимо обязательно самостоятельно изучать соответствующий материал в Moodle, фиксируя записи в тетради, а также выполнять практические задания.

Подготовка к зачету с оценкой осуществляется по представленным в списке основной и дополнительной литературе, а также частично по нормативным документам. В учебниках и учебных пособиях содержатся одноименные параграфы, что позволит успешно подготовиться к зачету с оценкой. Рекомендуемые литература и интернет-ресурсы будут полезны при выполнении практических заданий, при чтении кода программ и для подготовки к тестированиям. В рекомендуемых интернет-ресурсах также можно найти ссылки на научные журналы по информационным технологиям и на форумы профессиональных сообществ программистов.

Методические рекомендации по составлению конспекта

Конспект — сложный способ изложения содержания книги или статьи в логической последовательности. Внимательно прочитайте текст. Уточните в справочной литературе непонятные слова. При записи не забудьте вынести справочные данные на поля конспекта. Выделите главное, составьте план, представляющий собой перечень заголовков, подзаголовков, вопросов, последовательно раскрываемых затем в конспекте.

Законспектируйте материал, четко следуя пунктам плана. При конспектировании старайтесь выразить мысль своими словами. Записи следует вести четко, ясно.

При оформлении конспекта необходимо стремиться к емкости каждого предложения.

Методические рекомендации по составлению опорного конспекта

Опорный конспект — вид внеаудиторной самостоятельной работы студента по созданию краткой информационной структуры, обобщающей и отражающей суть материала лекции, темы учебника.

Опорный конспект — это наилучшая форма подготовки к ответу на вопросы.

Основная цель опорного конспекта — облегчить запоминание. Этапы составления опорного конспекта:

1. Изучить материалы темы, выбрать главное и второстепенное;
2. Установить логическую связь между элементами темы;
3. Представить характеристику элементов в краткой форме;
4. Выбрать опорные сигналы для акцентирования главной информации и отобразить в структуре работы.

Методические рекомендации по прохождению тестирования

Тестирование — это исследовательский метод, который позволяет выявить уровень знаний, умений и навыков, способностей, а также их соответствие определенным нормам усвоения, путем выполнения испытуемым ряда специальных заданий.

Следует понимать, что тестовые задания могут быть представлены в различных формах:

— задания закрытой формы, в которых обучающийся выбирает один или несколько правильных ответов из заданного набора:

— задания на дополнение (открытые задания) требующие самостоятельного получения ответов:

— задания на установления соответствия (с множественным выбором), выполнение которых связано с выявлением соответствия между элементами нескольких множеств:

— задания на установление правильной последовательности, в которых от учащегося требует указать порядок действий или процессов и другие. Этапы подготовки к тестированию:

1. Внимательно прочитайте материал по конспекту, составленному на учебном занятии. Прочитайте тот же материал по учебнику, учебному пособию.

2. Постарайтесь разобраться с непонятным, в частности новыми терминами и конструкциями.

3. Ответьте на контрольные вопросы для самопроверки, имеющиеся в учебнике, конспекте и т. д.

4. Кратко перескажите содержание изученного материала «своими словами».

5. Выучите определения основных понятий, условные обозначения, формулы и конструкции.

Подготовка к практическим занятиям

В ходе подготовки к практическим занятиям необходимо изучить основную литературу, ознакомиться с дополнительной литературой, новыми публикациями в периодических изданиях, ознакомиться с программным обеспечением. Следует дорабатывать свой конспект лекции, делая в нем соответствующие записи из литературы, рекомендованной преподавателем и предусмотренной учебной программой.

Заканчивать подготовку следует закреплением материала с использованием соответствующих программных продуктов.

Все практические задания, предусмотренные рабочей программой, представлены в фонде оценочных средств по дисциплине.

Критерии оценивания выполненных практических работ:

- правильность выполнения работы (отсутствие фактических, логических и других ошибок);
- полнота выполнения работы;
- своевременность выполнения;
- правильность оформления отчета.

За задания, выполненные позже установленного срока или с нарушениями требований к оформлению, оценка на балл снижается.

Порядок организации самостоятельной работы студентов

Самостоятельная работа студентов способствует развитию самостоятельности, ответственности и организованности, творческого подхода к решению проблем учебного и профессионального уровня.

Целью самостоятельной работы студентов является: овладение практическими знаниями, профессиональными умениями и навыками деятельности по специальности, опытом творческой, исследовательской деятельности.

Самостоятельная работа студентов предполагает:

- самостоятельный поиск ответов и необходимой информации в рамках изучаемых тем;
- выполнение заданий для самостоятельной работы, в том числе тестов;
- изучение теоретического и лекционного материала, а также основной и дополнительной литературы при подготовке к практическим занятиям.

Самостоятельная работа студентов является обязательным элементом подготовки специалиста среднего звена.

5 Учебная литература и ресурсы информационно-телекоммуникационной сети «Интернет»

Основная литература

1. Исаченко, О. В. Базы данных : учебное пособие / О. В. Исаченко. — Москва : ИНФРА-М, 2025. — 202 с. — (Среднее профессиональное образование). - ISBN 978-5-16-016506-6. - Текст : электронный. - URL: <https://znanium.ru/catalog/product/1171948> (дата обращения: 04.08.2026).

2. Шустова, Л. И. Базы данных : учебник / Л.И. Шустова, О.В. Тараканов. — Москва : ИНФРА-М, 2026. — 304 с. + Доп. материалы [Электронный ресурс]. — (Среднее профессиональное образование). - ISBN 978-5-16-014161-9. - Текст : электронный. - URL: <https://znanium.ru/catalog/product/2216840> (дата обращения: 04.08.2026).

Дополнительная литература

1. Анализ данных: учебник для вузов / под редакцией В. С. Мхитаряна. — Москва : Издательство Юрайт, 2025. — 448 с. — (Высшее образование). — ISBN 978-5-534-19964-2. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/560311> (дата обращения: 15.12.2025).

2. Колошкина, И. Е. Автоматизация проектирования технологической документации: учебник и практикум для среднего профессионального образования / И. Е. Колошкина. — Москва : Издательство Юрайт, 2025. — 371 с. — (Профессиональное образование). — ISBN 978-5-534-13635-7. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/567275> (дата обращения: 15.12.2025).

3. Тружкин, Н. П. Базы данных: проектирование: учебник для среднего профессионального образования / Н. П. Стружкин, В. В. Годин. — Москва : Издательство Юрайт, 2025. — 477 с. — (Профессиональное образование). — ISBN 978-5-534-11635-9. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/566509> (дата обращения: 15.12.2025).

Нормативные документы

1. ГОСТ 12207 ИТ. Процессы жизненного цикла программных средств. — URL: <https://normativ.kontur.ru/document?moduleId=9&documentId=212198&ysclid=mhgyd44gx109597256>

2. ГОСТ 19.101-2024. Единая система программной документации. Виды программ и программных документов. (введен в действие 30.01.2025). — URL: <https://protect.gost.ru/document.aspx?control=7&id=263735&ysclid=mhgygb9erh989749121>

3. ГОСТ 19.101-77. Единая система программной документации (ЕСПД). — URL: https://rosgosts.ru/file/gost/35/080/gost_19.101-77.pdf

4. ГОСТ 19.201-78 Техническое задание, требования к содержанию и оформлению. — URL: https://rosgosts.ru/file/gost/35/080/gost_19.201-78.pdf

5. ГОСТ 19.401-78. Единая система программной документации. Текст программы. Требования к содержанию и оформлению. — URL: <https://docs.cntd.ru/document/1200007651?ysclid=mhqzq4l03u398598150>

6. ГОСТ Р ИСО/МЭК 15910-2002 Информационная технология (ИТ). Процесс создания документации пользователя программного средства. — URL: <https://docs.cntd.ru/document/1200030141?ysclid=mhqzifmkdo827279089>
7. Стандарты ISO C++. — URL: <https://isocpp.org/std/the-standard>

Интернет-ресурсы

1. GitFlic. Российская платформа для работы с кодом. — URL: <https://gitflic.ru/>
2. Online Python IDE. — URL: <https://www.online-python.com/>
3. Stack Overflow (на русском). — URL: <https://ru.stackoverflow.com/?ysclid=mhqyx9542h717626304>
4. The Linux Kernel documentation. — URL: <https://www.kernel.org/doc/html/latest/>
5. SWEBOK: Руководство по основам программной инженерии. — URL: <https://www.computer.org/education/bodies-of-knowledge/software-engineering>
6. Документация по разработке приложений для Astra Linux. — URL: <https://docs.astralinux.ru/latest/?ysclid=mhqz4o3x8q692967251>
7. Документация по разработке приложений для Windows. — URL: <https://learn.microsoft.com/ru-ru/windows/apps/>
8. Научный журнал «Прикладная информатика» (Московский финансово-промышленный университет «Синергия»). — URL: https://www.elibrary.ru/title_about_new.asp?id=25599
9. Научный журнал «Программная инженерия» (ООО «Издательство «Новые технологии»). — URL: https://www.elibrary.ru/title_about_new.asp?id=32250
10. Реестр российского программного обеспечения. — URL: <https://reestr.digital.gov.ru/>
11. Роспатент. Федеральная служба по интеллектуальной собственности. — URL: <https://rospatent.gov.ru/ru>
12. РОССТАНДАРТ. Федеральное агентство по техническому регулированию и метрологии. — URL: <https://www.rst.gov.ru/portal/gost>

6 Материально-техническая база, информационные технологии, программное обеспечение и информационные справочные системы

Для реализации дисциплины необходимы:

Лаборатория программирования и баз данных, включающая:

- компьютерный класс (15-20 рабочих мест) с современными ПК, объединенными в локальную сеть с выходом в Интернет;
- проектор и экран;
- маркерная или меловая доска;
- информационно-коммуникационные технологии;
- программное обеспечение.

Информационно-коммуникационные технологии

Локальная вычислительная сеть с организованным доступом к электронным образовательным ресурсам.

Система видеоконференцсвязи (Mts Link) для проведения дистанционных консультаций.

Виртуальная образовательная среда на базе LMS (Moodle) для размещения учебных материалов и проведения тестирования.

Система облачного хранения (Яндекс Диск) и системы для коллективной работы над проектами (Яндекс Документы, Mts Link.Доски).

Программное обеспечение

Windows 10/11, Linux (дистрибутивы Debian, Ubuntu, Astra или Alt), MS Visual Studio, MS Visual Studio Code, SQLite, PostgreSQL, pgAdmin, DBeaver, Python IDE, 1С:Предприятие, Eclipse IDE, Apache NetBeans, Spacemacs, AndroidStudio, Draw.io, StarUML 5, Inkscape, LibreOffice, Oracle VM VirtualBox, Notepad++.

Электронно-библиотечные системы (ЭБС)

1. Электронно-библиотечная система «IPR SMART» (СПО) — URL: <https://www.iprbookshop.ru/>;
2. Электронно-библиотечная система «Лань» — URL: <https://e.lanbook.com/>;
3. Электронно-библиотечная система «iBooks» — URL: <https://ibooks.ru/>;
4. Электронно-библиотечная система «ZNANIUM» — URL: <https://znanium.ru/>.