

Документ подписан электронной подписью

Информация о владельце:

ФИО: Выборнова Любовь Алексеевна

Должность: Ректор

Дата подписания: 31.08.2026 09:48:13

Уникальный программный ключ:

c3b3b9c625f6c113afa2a2c42baff9e05a38b76e

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное бюджетное образовательное учреждение высшего образования
«Поволжский государственный университет сервиса» (ФГБОУ ВО «ПВГУС»)

Высшая школа интеллектуальных систем и кибертехнологий

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

«АНАЛИЗ ДАННЫХ»

Программа бакалавриата по направлению подготовки

09.03.03 «Прикладная информатика»

Направленность (профиль):

«Цифровые решения для бизнеса»

Квалификация выпускника: **бакалавр**

Рабочая программа дисциплины «Анализ данных» разработана в соответствии с федеральным государственным образовательным стандартом высшего образования - *бакалавриат* по направлению подготовки 09.03.03 «Прикладная информатика», утвержденным приказом Министерства науки и высшего образования Российской Федерации от 17 ноября 2020 г. N 1427.

Составители:

старший преподаватель

(ученая степень, ученое звание)

Коновалова А.Д.

(ФИО)

РПД обсуждена на заседании Высшей школы интеллектуальных систем и кибертехнологий

« » _____ 202_ г., протокол № 2 _____

Директор ВШИСиКТ

к.э.н., доцент

(уч.степень, уч.звание)

Филлипова О.А.

(ФИО)

1. ПЕРЕЧЕНЬ ПЛАНИРУЕМЫХ РЕЗУЛЬТАТОВ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ, СООТНЕСЕННЫХ С ПЛАНИРУЕМЫМИ РЕЗУЛЬТАТАМИ ОСВОЕНИЯ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ

1.1. Цель освоения дисциплины

Целью освоения дисциплины является:

- формирование у обучающихся общепрофессиональной компетенции, обеспечивающей анализ организационно-технических и экономических процессов с применением методов системного анализа, статистического анализа и математического моделирования данных;
- формирование профессиональной компетенции по разработке структуры баз данных информационных систем в соответствии с архитектурной спецификацией и задачами аналитической обработки данных.

1.2. Перечень планируемых результатов обучения по дисциплине

Код и наименование компетенции	Код и наименование индикатора достижения компетенции	Планируемые результаты обучения по дисциплине	Основание (ПС) *для профессиональных компетенций
ОПК-6. Способен анализировать и разрабатывать организационно-технические и экономические процессы с применением методов системного анализа и математического моделирования	ИОПК-6.1. Применяет знания основ теории систем и системного анализа, дискретной математики, математической логики, теории алгоритмов, математического и имитационного моделирования для автоматизации задач принятия решений, анализа информационных потоков, расчета экономической эффективности и надежности информационных систем и технологий. ИОПК-6.2. Проводит инженерные расчеты основных показателей результативности создания и применения информационных систем и технологий.	Знает: основы системного подхода к анализу данных; методы описательной статистики, статистического вывода, математического моделирования и оценки качества моделей; показатели результативности информационных систем и технологий, применяемые при аналитическом обосновании решений. Умеет: формализовать задачу анализа данных, исследовать и подготавливать данные, выбирать и применять статистические и математические методы, рассчитывать и интерпретировать показатели результативности, обосновывать выводы с учетом ограничений исходных данных и моделей. Владеет: навыками разведочного и статистического анализа данных, построения и оценки аналитических моделей, расчета показателей, документирования и интерпретации результатов для поддержки принятия решений.	–
ПК-5. Способен к выполнению работ по управлению, настройке, разработке баз данных ИС	ИПК-5.1. Осуществляет разработку структуры баз данных ИС в соответствии с архитектурной спецификацией	Знает: принципы проектирования структур данных и реляционных баз данных, сущности, атрибуты, ключи и ограничения целостности, нормализацию, основы аналитических схем и требования архитектурной спецификации к структуре БД. Умеет: проектировать логическую структуру базы данных и аналитической витрины по архитектурной спецификации, определять сущности, атрибуты и связи, формировать структуру таблиц и запросы для получения аналитических показателей. Владеет: навыками разработки и реализации фрагментов структуры аналитической базы данных, подготовки SQL-запросов к данным и согласования структуры хранения с задачами анализа и архитектурной спецификацией ИС.	–

2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ

Дисциплина относится к Блоку 1 «Дисциплины (модули)» образовательной программы (Модуль специализации (по выбору). Специализация 3 «Системная и бизнес аналитика»).

3. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

3.1. Объем и структура дисциплины

Общая трудоемкость дисциплины составляет 5 з.е. (180 час.), их распределение по видам работ для очной формы обучения представлено в таблице.

Виды учебных занятий и работы обучающихся	Трудоемкость, час
Общая трудоемкость дисциплины, час	180
Контактная работа обучающихся с преподавателем по видам учебных занятий (всего), в т.ч.:	48
занятия лекционного типа (лекции)	18
занятия семинарского типа (семинары, практические занятия, практикумы, коллоквиумы и иные аналогичные занятия)	18
лабораторные работы	12
Самостоятельная работа всего, в т.ч.:	105
Самоподготовка по темам (разделам) дисциплины	105
Контроль (часы на экзамен)	27
Промежуточная аттестация	Экзамен

Примечание: -/- объем часов соответственно для очной, очно-заочной форм обучения

Дисциплина реализуется посредством проведения контактной работы с обучающимися (включая проведение текущего контроля успеваемости), самостоятельной работы обучающихся и промежуточной аттестации.

В процессе освоения дисциплины может применяться электронное обучение и дистанционные образовательные технологии.

В процессе освоения дисциплины обучающиеся обеспечены доступом к электронной информационно-образовательной среде и электронно-библиотечным системам.

3.2. Содержание дисциплины, структурированное по темам

Планируемые результаты освоения: код формируемой компетенции и индикаторы достижения компетенций	Наименование разделов, тем	Виды учебной работы				Формы текущего контроля (наименование оценочного средства)
		Контактная работа			Самостоятельная работа, час	
		Лекции, час	Лабораторные работы, час	Практические занятия, час		
ОПК-6: ИОПК-6.1 ИОПК-6.2 ПК-5: ИПК-5.1	ТЕМА 1. ВВЕДЕНИЕ В АНАЛИЗ ДАННЫХ. ЖИЗНЕННЫЙ ЦИКЛ АНАЛИТИЧЕСКОГО ИССЛЕДОВАНИЯ Содержание лекции: 1. Место анализа данных в прикладной информатике, системной и бизнес-аналитике. Цели описательного, диагностического, прогностического и предписывающего анализа. 2. Типы данных, шкалы измерения, источники данных и требования к их качеству. Понятия выборки, признака, целевой переменной и показателя. 3. Этапы аналитического исследования: постановка задачи, сбор и подготовка данных, анализ и моделирование, оценка, интерпретация и представление результатов. Воспроизводимость аналитического процесса.	2				Лекция-визуализация (в т.ч. в ЭИОС) Тестирование по теме лекционного занятия
	Лабораторная работа 1. Подготовка аналитической среды. Загрузка табличного набора данных в Python/Jupyter, исследование структуры, типов полей и базовых характеристик.		2			Отчет по лабораторной работе
	Практическая работа 1. Формализация аналитической задачи: определение объекта анализа, информационных потоков, показателей результативности, состава данных и критериев результата.			2		Отчет по практической работе
	Самостоятельная работа.				11	Самостоятельное изучение учебных материалов

Планируемые результаты освоения: код формируемой компетенции и индикаторы достижения компетенций	Наименование разделов, тем	Виды учебной работы				Формы текущего контроля (наименование оценочного средства)
		Контактная работа			Самостоятельная работа, час	
		Лекции, час	Лабораторные работы, час	Практические занятия, час		
ОПК-6: ИОПК-6.1 ИОПК-6.2 ПК-5: ИПК-5.1	ТЕМА 2. ИСТОЧНИКИ ДАННЫХ И ПОДГОТОВКА ДАННЫХ К АНАЛИЗУ Содержание лекции: 1. Структурированные, полуструктурированные и неструктурированные источники. Табличные файлы, реляционные БД, выгрузки информационных систем. 2. Профилирование качества: пропуски, дубликаты, противоречия, аномальные значения, неверные типы и форматы. Правила очистки и документирование преобразований. 3. Преобразование признаков: кодирование категорий, масштабирование, нормализация, агрегирование. Риски утечки данных и искажения аналитических выводов.	2				Лекция-визуализация (в т.ч. в ЭИОС) Тестирование по теме лекционного занятия
	Лабораторная работа 2. Очистка набора данных: обработка пропусков и дубликатов, преобразование типов, контроль диапазонов, подготовка воспроизводимого конвейера предобработки.		2			Отчет по лабораторной работе
	Практическая работа 2. Построение профиля качества данных и разработка обоснованной стратегии очистки для данных информационной системы.			2		Отчет по практической работе
	Самостоятельная работа.				12	Самостоятельное изучение учебных материалов

Планируемые результаты освоения: код формируемой компетенции и индикаторы достижения компетенций	Наименование разделов, тем	Виды учебной работы				Формы текущего контроля (наименование оценочного средства)
		Контактная работа			Самостоятельная работа, час	
		Лекции, час	Лабораторные работы, час	Практические занятия, час		
ОПК-6: ИОПК-6.1 ИОПК-6.2 ПК-5: ИПК-5.1	ТЕМА 3. ОПИСАТЕЛЬНАЯ СТАТИСТИКА И РАЗВЕДОЧНЫЙ АНАЛИЗ ДАННЫХ Содержание лекции: 1. Меры центральной тенденции и положения: среднее, медиана, мода, квантили. Меры разброса: размах, дисперсия, стандартное отклонение, межквартильный размах. 2. Распределения и их графическое представление. Анализ выбросов. Ковариация и корреляция; ограничения корреляционного анализа. 3. Разведочный анализ данных (EDA): последовательность действий, формирование гипотез, расчет и интерпретация показателей результативности информационных процессов.	2				Лекция- визуализация (в т.ч. в ЭИОС) Тестирование по теме лекционного занятия
	Лабораторная работа 3. Выполнение разведочного анализа: описательные статистики, распределения, выбросы, взаимосвязи признаков и формулирование аналитических выводов.		2			Отчет по лабораторной работе
	Практическая работа 3. Расчет и интерпретация ключевых показателей результативности по данным бизнес-процесса или информационной системы.			2		Отчет по практической работе
	Самостоятельная работа.				12	Самостоятельное изучение учебных материалов

Планируемые результаты освоения: код формируемой компетенции и индикаторы достижения компетенций	Наименование разделов, тем	Виды учебной работы				Формы текущего контроля (наименование оценочного средства)
		Контактная работа			Самостоятельная работа, час	
		Лекции, час	Лабораторные работы, час	Практические занятия, час		
ОПК-6: ИОПК-6.1 ИОПК-6.2 ПК-5: ИПК-5.1	ТЕМА 4. МОДЕЛИРОВАНИЕ ДАННЫХ И СТРУКТУРА АНАЛИТИЧЕСКОЙ БАЗЫ ДАННЫХ Содержание лекции: 1. Реляционная модель данных: сущности, атрибуты, домены, первичные и внешние ключи, ограничения целостности, типы связей. 2. Нормализация данных и функциональные зависимости. Назначение нормальных форм при проектировании структуры БД. 3. Структуры данных для аналитики: операционные и аналитические модели, таблицы фактов и измерений, аналитические витрины. Связь структуры БД с архитектурной спецификацией ИС.	2				Лекция-визуализация (в т.ч. в ЭИОС) Тестирование по теме лекционного занятия
	Лабораторная работа 4. Реализация структуры аналитической базы данных в PostgreSQL/SQLite: создание таблиц, ключей, связей и ограничений по заданной спецификации.		2			Отчет по лабораторной работе
	Практическая работа 4. Проектирование логической модели данных и аналитической витрины по архитектурной спецификации: сущности, атрибуты, связи, ключи и ограничения.			2		Отчет по практической работе
	Самостоятельная работа.				12	Самостоятельное изучение учебных материалов

Планируемые результаты освоения: код формируемой компетенции и индикаторы достижения компетенций	Наименование разделов, тем	Виды учебной работы				Формы текущего контроля (наименование оценочного средства)
		Контактная работа			Самостоятельная работа, час	
		Лекции, час	Лабораторные работы, час	Практические занятия, час		
ОПК-6: ИОПК-6.1 ИОПК-6.2 ПК-5: ИПК-5.1	ТЕМА 5. SQL ДЛЯ АНАЛИЗА ДАННЫХ И АГРЕГИРОВАНИЕ ПОКАЗАТЕЛЕЙ Содержание лекции: 1. Выборка, фильтрация и сортировка данных. Соединения таблиц и объединение информационных потоков. 2. Агрегатные функции, GROUP BY и HAVING. Подзапросы, общие табличные выражения. Принципы расчета показателей на уровне БД. 3. Оконные функции, ранжирование, накопительные показатели и сравнение периодов. Организация аналитических запросов к данным ИС.	2				Лекция-визуализация (в т.ч. в ЭИОС) Тестирование по теме лекционного занятия
	Лабораторная работа 5. Разработка аналитических SQL-запросов: соединения, группировки, CTE и оконные функции для расчета показателей.		2			Отчет по лабораторной работе
	Практическая работа 5. Проектирование набора SQL-запросов для анализа информационных потоков и расчета показателей результативности.			2		Отчет по практической работе
	Самостоятельная работа.				12	Самостоятельное изучение учебных материалов

Планируемые результаты освоения: код формируемой компетенции и индикаторы достижения компетенций	Наименование разделов, тем	Виды учебной работы				Формы текущего контроля (наименование оценочного средства)
		Контактная работа			Самостоятельная работа, час	
		Лекции, час	Лабораторные работы, час	Практические занятия, час		
ОПК-6: ИОПК-6.1 ИОПК-6.2 ПК-5: ИПК-5.1	ТЕМА 6. СТАТИСТИЧЕСКИЙ ВЫВОД И ПРОВЕРКА ГИПОТЕЗ Содержание лекции: 1. Генеральная совокупность и выборка. Выборочные оценки и их свойства. Доверительные интервалы и уровень доверия. 2. Нулевая и альтернативная гипотезы, уровень значимости, р-значение. Ошибки первого и второго рода, статистическая мощность. 3. Выбор параметрических и непараметрических критериев. Сравнение групп и показателей; принципы анализа результатов экспериментов и А/В-сравнений.	2				Лекция-визуализация (в т.ч. в ЭИОС) Тестирование по теме лекционного занятия
	Лабораторная работа 6. Проверка статистических гипотез в Python: выбор критерия, расчет статистики и р-значения, интерпретация результата и ограничений.		2			Отчет по лабораторной работе
	Практическая работа 6. Выбор метода статистического сравнения и подготовка аналитического заключения по результатам исследования.			2		Отчет по практической работе
	Самостоятельная работа.				12	Самостоятельное изучение учебных материалов

Планируемые результаты освоения: код формируемой компетенции и индикаторы достижения компетенций	Наименование разделов, тем	Виды учебной работы				Формы текущего контроля (наименование оценочного средства)
		Контактная работа			Самостоятельная работа, час	
		Лекции, час	Лабораторные работы, час	Практические занятия, час		
ОПК-6: ИОПК-6.1 ИОПК-6.2 ПК-5: ИПК-5.1	ТЕМА 7. РЕГРЕССИОННЫЙ АНАЛИЗ И ПРОГНОЗИРОВАНИЕ Содержание лекции: 1. Задачи регрессии. Простая и множественная линейная регрессия. Интерпретация коэффициентов и предпосылки модели. 2. Качество регрессионной модели: MAE, MSE, RMSE, коэффициент детерминации. Анализ остатков и устойчивость выводов. 3. Разделение данных на обучающую и тестовую части, кросс-валидация. Базовые подходы к прогнозированию показателей и ограничения экстраполяции.	2				Лекция- визуализация (в т.ч. в ЭИОС) Тестирование по теме лекционного занятия
	Практическая работа 7. Построение и интерпретация регрессионной модели для прогнозирования показателя; сравнение качества альтернативных моделей.			2		Отчет по практической работе
	Самостоятельная работа.				12	Самостоятельное изучение учебных материалов

Планируемые результаты освоения: код формируемой компетенции и индикаторы достижения компетенций	Наименование разделов, тем	Виды учебной работы				Формы текущего контроля (наименование оценочного средства)
		Контактная работа			Самостоятельная работа, час	
		Лекции, час	Лабораторные работы, час	Практические занятия, час		
ОПК-6: ИОПК-6.1 ИОПК-6.2 ПК-5: ИПК-5.1	ТЕМА 8. КЛАССИФИКАЦИЯ, СЕГМЕНТАЦИЯ И ОЦЕНКА АНАЛИТИЧЕСКИХ МОДЕЛЕЙ Содержание лекции: 1. Постановка задачи классификации. Базовые модели: логистическая регрессия, деревья решений. Матрица ошибок и метрики accuracy, precision, recall, F1, ROC-AUC. 2. Постановка задачи кластеризации. Понятие расстояния и сходства, алгоритм k-средних, интерпретация сегментов. 3. Переобучение и недообучение. Валидация, выбор метрики, интерпретируемость моделей и ответственность при принятии решений на основе данных.	2				Лекция- визуализация (в т.ч. в ЭИОС) Тестирование по теме лекционного занятия
	Практическая работа 8. Сравнение результатов классификации или сегментации: расчет метрик, выбор модели и обоснование управленческой интерпретации.			2		Отчет по практической работе
	Самостоятельная работа.					11

Планируемые результаты освоения: код формируемой компетенции и индикаторы достижения компетенций	Наименование разделов, тем	Виды учебной работы				Формы текущего контроля (наименование оценочного средства)
		Контактная работа			Самостоятельная работа, час	
		Лекции, час	Лабораторные работы, час	Практические занятия, час		
ОПК-6: ИОПК-6.1 ИОПК-6.2 ПК-5: ИПК-5.1	ТЕМА 9. КОМПЛЕКСНЫЙ АНАЛИТИЧЕСКИЙ ПРОЕКТ И ИНТЕРПРЕТАЦИЯ РЕЗУЛЬТАТОВ Содержание лекции: 1. Интеграция этапов анализа: получение данных, структура хранения, подготовка, исследование, моделирование и оценка. Прослеживаемость преобразований. 2. Документирование аналитического решения: постановка задачи, описание данных и структуры БД, методы, показатели качества, ограничения и воспроизводимость. 3. Представление результатов для заинтересованных сторон. Этические аспекты, конфиденциальность, качество исходных данных и границы применимости аналитических выводов.	2				Лекция-визуализация (в т.ч. в ЭИОС) Тестирование по теме лекционного занятия
	Практическая работа 9. Проектирование сквозного аналитического решения: архитектура данных, структура БД, методы анализа, показатели качества и форма представления результата.			2		Отчет по практической работе
	Самостоятельная работа.				11	Самостоятельное изучение учебных материалов
ИТОГО		18	12	18	105	

Примечание: -/- объем часов соответственно для очной, очно-заочной форм обучения

4. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ

4.1. Общие методические рекомендации по освоению дисциплины, образовательные технологии

Дисциплина реализуется посредством проведения контактной работы с обучающимися (включая проведение текущего контроля успеваемости), самостоятельной работы обучающихся и промежуточной аттестации.

При проведении учебных занятий по дисциплине обеспечивается развитие у обучающихся навыков командной работы, межличностной коммуникации, принятия решений, лидерских качеств (включая проведение интерактивных лекций, групповых дискуссий, ролевых игр, тренингов, анализ ситуаций и имитационных моделей, преподавание дисциплины в форме курса, составленного на основе результатов научных исследований, проводимых университетом, в том числе с учетом региональных особенностей профессиональной деятельности выпускников и потребностей работодателей).

Преподавание дисциплины ведется с применением следующих видов образовательных технологий:

- балльно-рейтинговая технология оценивания;
- электронное обучение;

Для оценки знаний, умений, навыков и уровня сформированности компетенции по дисциплине применяется балльно-рейтинговая система контроля и оценки успеваемости студентов. В основу балльно-рейтинговой системы положены принципы, в соответствии с которыми формирование рейтинга студента осуществляется в ходе текущего контроля успеваемости. Максимальное количество баллов в семестре – 100.

4.2. Методические указания для обучающихся по освоению дисциплины на занятиях лекционного типа

Лекционный курс предполагает систематизированное изложение математических и инженерных основ функционального программирования. При разборе каждой темы теоретические положения сопровождаются короткими фрагментами F#, схемами вычислений и сравнением с императивным стилем. Особое внимание уделяется пониманию семантики чистых функций, типов, композиции и управления эффектами.

Лекционные занятия проводятся в поточной аудитории с применением мультимедийного проектора в виде учебной презентации или в ЭИОС университета.

В ходе лекционных занятий рекомендуется конспектирование учебного материала.

Возможно ведение конспекта лекций в виде интеллект-карт.

Отдельные темы предлагаются для самостоятельного изучения (конспектируются).

Материалы лекций являются опорной основой для подготовки обучающихся к практическим занятиям / лабораторным работам и выполнения заданий самостоятельной работы, а также к мероприятиям текущего контроля успеваемости и промежуточной аттестации по дисциплине.

4.3. Методические указания для обучающихся по освоению дисциплины на лабораторных работах

Подготовку к каждой лабораторной работе следует начинать с изучения теоретического материала и сигнатур используемых функций и типов. В ходе работы обучающийся реализует программное решение на F#, проверяет его на тестовых данных и оформляет отчет, содержащий постановку задачи, исходный код ключевых функций, результаты тестирования и выводы. Каждая работа подлежит защите преподавателю.

При оценивании лабораторных работ учитывается следующее:

- качество выполнения экспериментально-практической части работы и степень соответствия результатов работы заданным требованиям;
- качество оформления отчета по работе;
- качество устных ответов на контрольные вопросы при защите работы.

Лабораторные работы организуются, в том числе в форме практической подготовки, которая предусматривает участие обучающихся в выполнении отдельных элементов работ, связанных с будущей профессиональной деятельностью.

4.4. Методические указания для обучающихся по освоению дисциплины на занятиях семинарского типа/ на практических занятиях

Практические занятия направлены на формализацию прикладных задач, функциональную декомпозицию, выбор типов данных и способов композиции функций. Основной формой работы являются разбор предметных ситуаций, проектирование сигнатур и моделей данных, сравнительный анализ решений и защита принятых проектных решений.

Практические (семинарские) занятия обучающихся обеспечивают:

- проверку и уточнение знаний, полученных на лекциях;
- получение умений и навыков составления докладов и сообщений, обсуждения вопросов по учебному материалу дисциплины;
- подведение итогов занятий по рейтинговой системе, согласно технологической карте дисциплины.

Практические занятия организуются, в том числе в форме практической подготовки, которая предусматривает участие обучающихся в выполнении отдельных элементов работ, связанных с будущей профессиональной деятельностью.

4.5. Методические указания по самостоятельной работе обучающихся

Самостоятельная работа обеспечивает подготовку обучающегося к аудиторным занятиям и мероприятиям текущего контроля и промежуточной аттестации по изучаемой дисциплине. Результаты этой подготовки проявляются в активности обучающегося на занятиях и в качестве выполненных практических заданий и других форм текущего контроля.

Самостоятельная работа студентов включает:

1. Изучение учебной литературы по курсу.
2. Работу с ресурсами Интернет
3. Самостоятельное изучение учебных материалов

При выполнении заданий для самостоятельной работы рекомендуется проработка материалов лекций по каждой пройденной теме, а также изучение рекомендуемой литературы.

Самостоятельная работа является обязательной частью освоения дисциплины и направлена на углубление навыков функциональной декомпозиции, программирования и тестирования.

Для обеспечения самостоятельной работы обучающихся используется электронный учебный курс, созданный в ЭИОС университета <http://sdo.tolgas.ru/>.

5. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

5.1. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины

Вся литература, включенная в данный перечень, представлена в виде электронных ресурсов в электронной библиотеке университета (ЭБС). Литература, используемая в печатном виде, представлена в научной библиотеке университета в объеме не менее 0,25 экземпляров на одного обучающегося.

Основная литература

1. Хаттон, Г. Программирование на Haskell / Graham Hutton. – 2-е изд. – Cambridge University Press, 2016.
2. Bird, R. Thinking Functionally with Haskell / Richard Bird. – Cambridge University Press, 2015.
3. O’Sullivan, B., Goerzen, J., Stewart, D. Real World Haskell. – O’Reilly Media, 2008.

Дополнительная литература

4. Microsoft Learn. F# documentation: справочные и учебные материалы по языку F# и платформе .NET.
5. F# Software Foundation. Материалы по функциональному программированию и экосистеме F#.
6. Haskell 2010 Language Report. Официальное описание языка и стандартной библиотеки.

5.2. Профессиональные базы данных, информационно-справочные системы, интернет-ресурсы

1. КонсультантПлюс [Электронный ресурс]: справочная правовая система. – Режим доступа: <http://www.consultant.ru/>.
2. Электронная библиотечная система Поволжского государственного университета сервиса [Электронный ресурс]. – Режим доступа: <http://elib.tolgaz.ru/>.
3. Электронно-библиотечная система Znanium.com [Электронный ресурс]. – Режим доступа: <http://znanium.com/>.
4. Электронно-библиотечная система «Издательство Лань» [Электронный ресурс]. – Режим доступа: <https://e.lanbook.com/>.
5. Научная электронная библиотека eLIBRARY.RU [Электронный ресурс]. – Режим доступа: <http://elibrary.ru/defaultx.asp>.
6. Открытое образование [Электронный ресурс]. – Режим доступа: <https://openedu.ru/>.
7. Базы данных Всероссийского института научной и технической информации (ВИНИТИ РАН) по естественным, точным и техническим наукам [Электронный ресурс]. – Режим доступа: <http://www.viniti.ru>.

5.3. Программное обеспечение

Информационное обеспечение учебного процесса по дисциплине осуществляется с использованием следующего программного обеспечения (лицензионного и свободно распространяемого), в том числе отечественного производства:

Наименование программного обеспечения	Назначение	Условия доступа
Microsoft Windows	Операционная система	Лицензионное программное обеспечение университета
Microsoft Office / совместимый офисный пакет	Работа с табличными данными, текстовыми документами и презентациями	Лицензионное или свободно распространяемое программное обеспечение
СДО MOODLE	Электронное обучение, размещение материалов и контроль	Доступ через ЭИОС университета
Python 3.x	Язык и среда выполнения аналитического кода	Свободно распространяемое программное обеспечение
Jupyter Notebook / JupyterLab	Интерактивная среда анализа данных и документирования вычислений	Свободно распространяемое программное обеспечение
pandas, NumPy, SciPy, scikit-learn, matplotlib	Библиотеки обработки, статистического анализа, моделирования и визуализации данных	Свободно распространяемое программное обеспечение
PostgreSQL и/или SQLite	Система управления базами данных для лабораторных работ и аналитических запросов	Свободно распространяемое программное обеспечение

6. ОПИСАНИЕ МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЙ БАЗЫ, НЕОБХОДИМОЙ ДЛЯ ОСУЩЕСТВЛЕНИЯ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА ПО ДИСЦИПЛИНЕ

Помещения представляют собой учебные аудитории для проведения учебных занятий, предусмотренных учебным планом и рабочей программой дисциплины, оснащенные оборудованием и техническими средствами обучения.

Занятия лекционного типа. Учебные аудитории для занятий лекционного типа укомплектованы мебелью и техническими средствами обучения, служащими для представления учебной информации (стационарные или переносные наборы демонстрационного оборудования (проектор, экран, компьютер/ноутбук), учебно-наглядные пособия (презентации по темам лекций), обеспечивающие тематические иллюстрации, соответствующие данной программе дисциплины.

Занятия семинарского типа. Учебные аудитории для занятий семинарского типа укомплектованы мебелью и техническими средствами обучения, служащими для представления учебной информации (стационарные или переносные наборы демонстрационного оборудования (проектор, экран, компьютер/ноутбук). Для проведения лабораторных работ используется учебная аудитория «Лаборатория информационных технологий, информатики и методов программирования // Учебная аудитория для проведения занятий лекционного типа, семинарского типа, курсового проектирования, консультаций, текущего контроля и промежуточной аттестации // Помещение для хранения и профилактического обслуживания учебного оборудования (Г-413)», оснащенная следующим оборудованием: рабочие места обучающихся; Рабочее место преподавателя; Учебно-наглядные пособия, обеспечивающие тематические иллюстрации. Компьютер Асус Intel G530\iH61\GeForceGT430\4096 Mb DDR3\500Gb\LCD 22" - 14 шт., Компьютер в сборе iP2.0 LCD - 1 шт., принтер HP 1200 - 1 шт., мобильный экран проекционный DraperLuma - 1 шт., мобильный проектор Sanyo PLC, Сетевое оборудование D-Link DES-1024D, 24-портовый; локальная сеть 100 мб с доступом в Интернет 50 мб. Лаборатория оснащена кондиционером.

Лабораторные работы. Для проведения лабораторных работ используется учебная аудитория «Лаборатория информационных технологий, информатики и методов программирования // Учебная аудитория для проведения занятий лекционного типа, семинарского типа, курсового проектирования, консультаций, текущего контроля и промежуточной аттестации // Помещение для хранения и профилактического обслуживания учебного оборудования (Г-413)», оснащенная следующим оборудованием: рабочие места обучающихся; Рабочее место преподавателя; Учебно-наглядные пособия, обеспечивающие тематические иллюстрации. Компьютер Аксус Intel G530\iH61\GeForceGT430\4096 Mb DDR3\500Gb\LCD 22" - 14 шт., Компьютер в сборе iP2.0 LCD - 1 шт., принтер HP 1200 - 1 шт., мобильный экран проекционный DraperLuma - 1 шт., мобильный проектор Sanyo PLC, Сетевое оборудование D-Link DES-1024D, 24-портовый; локальная сеть 100 мб с доступом в Интернет 50 мб. Лаборатория оснащена кондиционером.

Промежуточная аттестация. Для проведения промежуточной аттестации по дисциплине используются компьютерные классы, оснащенные компьютерной техникой с возможностью подключения к сети Интернет и обеспечением доступа в электронную информационно-образовательную среду университета и/или учебные аудитории, укомплектованные мебелью и техническими средствами обучения.

Самостоятельная работа. Помещения для самостоятельной работы оснащены компьютерной техникой с возможностью подключения к сети «Интернет» и доступом к электронной информационно-образовательной среде университета. Для организации самостоятельной работы обучающихся используются:

- компьютерные классы университета;
- библиотека (медиазал), имеющая места для обучающихся, оснащенные компьютерами с доступом к базам данных и сети Интернет.

Электронная информационно-образовательная среда университета (ЭИОС). Каждый обучающийся в течение всего периода обучения обеспечен индивидуальным неограниченным доступом к электронной информационно-образовательной среде университета (ЭИОС) <http://sdo.tolgas.ru/> из любой точки, в которой имеется доступ к информационно-телекоммуникационной сети "Интернет", как на территории университета, так и вне ее.

ЭИОС университета обеспечивает:

- доступ к учебным планам, рабочим программам дисциплин (модулей), программам практик, электронным учебным изданиям и электронным образовательным ресурсам, указанным в рабочих программах дисциплин (модулей), программах практик;
- формирование электронного портфолио обучающегося, в том числе сохранение его работ и оценок за эти работы.

В случае реализации образовательной программы с применением электронного обучения, дистанционных образовательных технологий ЭИОС дополнительно обеспечивает:

фиксацию хода образовательного процесса, результатов промежуточной аттестации и результатов освоения образовательной программы;

проведение учебных занятий, процедур оценки результатов обучения, реализация которых предусмотрена с применением электронного обучения, дистанционных образовательных технологий;

взаимодействие между участниками образовательного процесса, в том числе синхронное и (или) асинхронное взаимодействия посредством сети "Интернет".

7. ОСОБЕННОСТИ ОРГАНИЗАЦИИ ОБУЧЕНИЯ ДЛЯ ЛИЦ С ОГРАНИЧЕННЫМИ ВОЗМОЖНОСТЯМИ ЗДОРОВЬЯ И ИНВАЛИДОВ

При необходимости рабочая программа дисциплины может быть адаптирована для обеспечения образовательного процесса инвалидов и лиц с ограниченными возможностями здоровья, в том числе для дистанционного обучения. Для этого требуется заявление студента (его законного представителя) и заключение психолого-медико-педагогической комиссии (ПМПК).

В случае необходимости, обучающимся из числа лиц с ограниченными возможностями здоровья (по заявлению обучающегося) а для инвалидов также в соответствии с индивидуальной программой реабилитации инвалида, могут предлагаться следующие варианты восприятия учебной информации с учетом их индивидуальных психофизических особенностей, в том числе с применением электронного обучения и дистанционных технологий:

- для лиц с нарушениями зрения: в печатной форме увеличенным шрифтом; в форме электронного документа; в форме аудиофайла (перевод учебных материалов в аудиоформат); в печатной форме на языке Брайля; индивидуальные консультации с привлечением тифлосурдопереводчика; индивидуальные задания и консультации.

- для лиц с нарушениями слуха: в печатной форме; в форме электронного документа; видеоматериалы с субтитрами; индивидуальные консультации с привлечением сурдопереводчика; индивидуальные задания и консультации.

- для лиц с нарушениями опорно-двигательного аппарата: в печатной форме; в форме электронного документа; в форме аудиофайла; индивидуальные задания и консультации.

8. ОЦЕНОЧНЫЕ МАТЕРИАЛЫ (ФОНД ОЦЕНОЧНЫХ СРЕДСТВ) ДЛЯ ТЕКУЩЕГО КОНТРОЛЯ УСПЕВАЕМОСТИ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ

8.1. Описание показателей и критериев оценивания компетенций на различных этапах их формирования, описание шкал оценивания

Для оценки знаний, умений, навыков и уровня сформированности компетенции по дисциплине применяется балльно-рейтинговая система контроля и оценки успеваемости студентов. В основу балльно-рейтинговой системы положены принципы, в соответствии с которыми формирование рейтинга студента осуществляется в ходе текущего контроля успеваемости. Максимальное количество баллов в семестре – 100.

Шкала оценки результатов освоения дисциплины, сформированности результатов обучения

Форма промежуточной аттестации	Уровень сформированности	100-балльная шкала, %	Шкала освоения, %	5-балльная дифференцированная оценка/балл
экзамен	допороговый	ниже 61	ниже 61	«неудовлетворительно» / 2
	пороговый	61–85,9	61–69,9	«удовлетворительно» / 3
	пороговый	61–85,9	70–85,9	«хорошо» / 4
	повышенный	86–100	86–100	«отлично» / 5

По итогам текущей успеваемости студенту может быть выставлена оценка по промежуточной аттестации в соответствии с набранными за семестр баллами (по накопительному рейтингу). Студентам, набравшим в ходе текущего контроля успеваемости по дисциплине от 61 до 100 баллов и выполнившим все обязательные виды запланированных учебных занятий, по решению преподавателя без прохождения промежуточной аттестации выставляется оценка в соответствии со шкалой оценки результатов освоения дисциплины.

Результат обучения считается сформированным (повышенный уровень), если теоретическое содержание курса освоено полностью; при устных собеседованиях студент исчерпывающе, последовательно, четко и логически стройно излагает учебный материал; свободно справляется с задачами, вопросами и другими видами заданий, требующих применения знаний, использует в ответе дополнительный материал; все предусмотренные рабочей учебной программой задания выполнены в соответствии с установленными требованиями, студент способен анализировать полученные результаты; проявляет самостоятельность при выполнении заданий, качество их выполнения оценено числом баллов от 86 до 100, что соответствует повышенному уровню сформированности результатов обучения.

Результат обучения считается сформированным (пороговый уровень), если теоретическое содержание курса освоено полностью; при устных собеседованиях студент последовательно, четко и логически стройно излагает учебный материал; справляется с задачами, вопросами и другими видами заданий, требующих применения знаний; все предусмотренные рабочей учебной программой задания выполнены в соответствии с установленными требованиями, студент способен анализировать полученные результаты; проявляет самостоятельность при выполнении заданий, качество их выполнения оценено числом баллов от 61 до 85,9, что соответствует пороговому уровню сформированности результатов обучения.

Результат обучения считается несформированным, если студент при выполнении заданий не демонстрирует знаний учебного материала, допускает ошибки, неуверенно, с большими затруднениями выполняет задания, не демонстрирует необходимых умений, качество выполненных заданий не соответствует установленным требованиям, качество их выполнения оценено числом баллов ниже 61, что соответствует допороговому уровню.

Формы текущего контроля успеваемости

Наименование оценочного средства	Количество контрольных мероприятий	Максимальный балл за одно мероприятие	Максимальное количество баллов
Посещение и работа на лекционных занятиях	9	2	18
Защита отчетов по лабораторным работам	6	6	36
Выполнение и защита практических работ	9	4	36
Творческое/аналитическое задание по интеграции материалов дисциплины	1	10	10
ИТОГО			100

Система оценивания представлена в электронном учебном курсе по дисциплине <http://sdo.tolgas.ru/>.

8.2. Типовые контрольные задания или иные материалы для проведения ТЕКУЩЕГО КОНТРОЛЯ УСПЕВАЕМОСТИ

8.2.1. Типовые задания к практическим занятиям

Практическая работа 1. Формализация аналитической задачи

Задание 1. Для заданного бизнес-процесса сформулировать цель анализа, объект, предмет и перечень вопросов, на которые должны ответить данные.

Задание 2. Определить не менее пяти измеримых показателей результативности и указать источники данных для их расчета.

Задание 3. Составить схему информационных потоков: источник → преобразование → показатель → пользователь результата.

Задание 4. Сформировать критерии достаточности данных и перечень рисков, способных исказить аналитический вывод.

Практическая работа 2. Профилирование качества и стратегия подготовки данных

Задание 1. По описанию набора данных выявить потенциальные проблемы качества: пропуски, дубликаты, нарушения формата, выбросы и противоречия.

Задание 2. Для каждого типа проблемы выбрать способ обработки и обосновать его с учетом цели исследования.

Задание 3. Разработать последовательность преобразований данных и определить контрольные проверки после каждого этапа.

Задание 4. Подготовить краткий протокол качества данных до и после обработки с указанием метрик полноты и согласованности.

Практическая работа 3. Описательная статистика и показатели результативности

Задание 1. Рассчитать для числовых признаков среднее, медиану, квантили, дисперсию, стандартное отклонение и межквартильный размах.

Задание 2. Сопоставить распределения показателей для двух групп и сформулировать содержательные выводы без подмены причинности корреляцией.

Задание 3. Рассчитать матрицу корреляций, определить сильные взаимосвязи и объяснить ограничения их интерпретации.

Задание 4. Рассчитать заданный показатель результативности информационного процесса и дать заключение о его динамике.

Практическая работа 4. Проектирование структуры аналитической базы данных

Задание 1. По архитектурной спецификации выделить сущности, атрибуты и связи предметной области.

Задание 2. Определить первичные и внешние ключи, ограничения целостности и обязательность атрибутов.

Задание 3. Нормализовать исходную структуру до третьей нормальной формы и обосновать выполненные преобразования.

Задание 4. Спроектировать аналитическую витрину или схему «звезда» для расчета заданных показателей и представить логическую модель.

Практическая работа 5. Аналитические запросы SQL

Задание 1. Составить запрос с соединением не менее трех таблиц для получения аналитической выборки.

Задание 2. Сформировать агрегатный запрос с GROUP BY и HAVING для расчета показателей по сегментам.

Задание 3. Разработать запрос с CTE для пошаговой подготовки набора данных к анализу.

Задание 4. Использовать оконную функцию для ранжирования или расчета накопительного показателя и объяснить полученный результат.

Практическая работа 6. Статистическая проверка гипотез

Задание 1. Сформулировать нулевую и альтернативную гипотезы для заданной прикладной ситуации.

Задание 2. Выбрать статистический критерий с учетом типа данных, независимости наблюдений и предпосылок применения.

Задание 3. Интерпретировать заданные значения статистики и р-значения при выбранном уровне значимости.

Задание 4. Объяснить риск ошибок первого и второго рода и сформулировать решение, не выходящее за пределы статистических данных.

Практическая работа 7. Регрессионный анализ

Задание 1. Выделить целевую переменную и объясняющие признаки для задачи прогнозирования показателя.

Задание 2. Интерпретировать коэффициенты линейной регрессии и определить направление влияния факторов.

Задание 3. Рассчитать или интерпретировать MAE, RMSE и R^2 для двух моделей и выбрать предпочтительную по заданному критерию.

Задание 4. По диагностике остатков определить возможные проблемы модели и предложить направления улучшения.

Практическая работа 8. Классификация и сегментация

Задание 1. По матрице ошибок рассчитать accuracy, precision, recall и F1 и объяснить различия между метриками.

Задание 2. Выбрать метрику качества классификации для ситуации с несбалансированными классами и обосновать выбор.

Задание 3. Интерпретировать результаты кластеризации и описать содержательный профиль каждого сегмента.

Задание 4. Сравнить две аналитические модели по качеству, устойчивости и интерпретируемости и сформулировать рекомендацию.

Практическая работа 9. Проектирование сквозного аналитического решения

Задание 1. Разработать схему аналитического процесса от источников данных до представления результата пользователю.

Задание 2. Определить структуру хранения данных и объяснить, как она поддерживает необходимые аналитические запросы.

Задание 3. Выбрать методы анализа и показатели качества для поставленной задачи, обосновать последовательность их применения.

Задание 4. Подготовить структуру итогового аналитического отчета: данные, методы, результаты, ограничения, выводы и рекомендации.

8.2.2. Типовые задания для лабораторных работ

Лабораторная работа 1. Подготовка аналитической среды и первичный обзор данных

Задание 1. Открыть Jupyter Notebook/JupyterLab, создать рабочий блокнот и зафиксировать версии используемых библиотек.

Задание 2. Загрузить табличный набор данных, вывести размерность, названия столбцов, типы данных и первые наблюдения.

Задание 3. Выполнить базовое профилирование: количество пропусков, уникальных значений и описательные характеристики.

Задание 4. Сохранить воспроизводимый отчет с кодом, результатами выполнения и краткими выводами о пригодности данных.

Лабораторная работа 2. Очистка и преобразование данных

Задание 1. Выявить и удалить полные дубликаты, задокументировать количество измененных записей.

Задание 2. Обработать пропуски выбранным способом, объяснить влияние решения на дальнейший анализ.

Задание 3. Преобразовать типы и форматы полей, проверить допустимые диапазоны и логические ограничения.

Задание 4. Сформировать функцию или последовательность операций предобработки и сравнить показатели качества данных до и после обработки.

Лабораторная работа 3. Разведочный анализ данных

Задание 1. Рассчитать описательные статистики для количественных и категориальных признаков.

Задание 2. Построить графические представления распределений и выявить потенциальные выбросы.

Задание 3. Исследовать взаимосвязи признаков с помощью корреляций и групповых агрегатов.

Задание 4. Подготовить не менее пяти содержательных выводов, каждый из которых должен опираться на рассчитанный показатель или визуализацию.

Лабораторная работа 4. Реализация структуры аналитической базы данных

Задание 1. Создать базу данных и таблицы в соответствии с разработанной логической моделью.

Задание 2. Задать первичные и внешние ключи, типы данных и ограничения целостности.

Задание 3. Загрузить тестовый набор данных и проверить корректность связей между таблицами.

Задание 4. Подготовить SQL-скрипт создания структуры БД и схему, показывающую соответствие архитектурной спецификации.

Лабораторная работа 5. Аналитические SQL-запросы

Задание 1. Выполнить выборку с JOIN для объединения фактов и справочников.

Задание 2. Рассчитать агрегированные показатели по группам и периодам с применением GROUP BY.

Задание 3. Использовать CTE для подготовки промежуточного аналитического набора данных.

Задание 4. Применить оконную функцию для ранжирования, скользящего или накопительного показателя и проверить результат.

Лабораторная работа 6. Проверка статистических гипотез

Задание 1. Подготовить данные для статистического сравнения и проверить основные предпосылки выбранного метода.

Задание 2. Рассчитать доверительный интервал для заданного показателя.

Задание 3. Выполнить проверку гипотезы в Python и получить статистику критерия и p-значение.

Задание 4. Сформулировать вывод на языке прикладной задачи, указав уровень значимости, ограничения и недопустимые причинные интерпретации.

8.3. Типовые контрольные задания или иные материалы для проведения ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ

Форма проведения промежуточной аттестации по дисциплине: экзамен (по результатам накопительного рейтинга или в форме компьютерного тестирования).

Устно-письменная форма по экзаменационным билетам предполагается, как правило, для сдачи академической задолженности.

Перечень вопросов и заданий для подготовки к экзамену

ОПК-6. Способен анализировать и разрабатывать организационно-технические и экономические процессы с применением методов системного анализа и математического моделирования

1. Понятие анализа данных и его место в прикладной информатике, системной и бизнес-аналитике.
2. Описательный, диагностический, прогнозный и предписывающий анализ: назначение и различия.
3. Этапы жизненного цикла аналитического исследования. Требования к воспроизводимости результата.
4. Типы данных и шкалы измерения. Влияние типа данных на выбор аналитического метода.
5. Источники данных информационных систем. Понятие информационного потока в аналитической задаче.
6. Критерии качества данных: полнота, согласованность, точность, актуальность, уникальность.
7. Пропуски в данных: причины, способы выявления и стратегии обработки.
8. Дубликаты, противоречивые записи и нарушения форматов: выявление и устранение.
9. Аномальные значения и выбросы: способы обнаружения и риски автоматического удаления.
10. Преобразование признаков: кодирование, масштабирование, нормализация и агрегирование.
11. Меры центральной тенденции и положения: среднее, медиана, мода, квантили.
12. Меры вариации: размах, дисперсия, стандартное отклонение, межквартильный размах.
13. Распределения данных и их графическое исследование. Разведочный анализ данных.
14. Ковариация и корреляция. Корреляционная связь и ограничения причинной интерпретации.
15. Понятие показателя результативности информационной системы или процесса. Принципы расчета и интерпретации.
16. Генеральная совокупность, выборка и репрезентативность. Выборочные оценки.
17. Доверительный интервал и уровень доверия: назначение и интерпретация.
18. Нулевая и альтернативная статистические гипотезы. Уровень значимости и р-значение.
19. Ошибки первого и второго рода. Статистическая мощность.
20. Параметрические и непараметрические критерии: условия выбора.
21. Принципы статистического сравнения двух групп и интерпретация результата.
22. А/В-сравнение: постановка, метрика, гипотеза, риски множественных проверок и преждевременного вывода.
23. Задача регрессии. Простая и множественная линейная регрессия.
24. Интерпретация коэффициентов регрессионной модели и ограничения этой интерпретации.
25. Метрики качества регрессии: MAE, MSE, RMSE, R^2 .
26. Анализ остатков и диагностика регрессионной модели.
27. Обучающая и тестовая выборки. Кросс-валидация и контроль переобучения.
28. Задача классификации. Логистическая регрессия и дерево решений как базовые модели.
29. Матрица ошибок. Метрики accuracy, precision, recall и F1.
30. ROC-кривая и ROC-AUC: назначение и ограничения.
31. Задача кластеризации. Понятие меры расстояния и алгоритм k-средних.
32. Интерпретация сегментов и оценка практической полезности кластеризации.
33. Переобучение и недообучение аналитической модели. Способы обнаружения и снижения риска.

34. Интерпретируемость моделей и требования к обоснованию аналитического решения.
35. Документирование аналитического исследования: данные, преобразования, методы, метрики, ограничения и выводы.

ПК-5. Способен к выполнению работ по управлению, настройке, разработке баз данных ИС

36. Реляционная модель данных: сущности, атрибуты, домены и отношения.
37. Первичные и внешние ключи. Ссылочная целостность и ограничения данных.
38. Типы связей между сущностями и способы их реализации в структуре БД.
39. Функциональные зависимости и назначение нормализации данных.
40. Первая, вторая и третья нормальные формы. Практический смысл нормализации.
41. Архитектурная спецификация ИС как основание для разработки структуры базы данных.
42. Логическое и физическое проектирование базы данных: назначение и результаты этапов.
43. Операционная и аналитическая модели данных: различия и области применения.
44. Таблица фактов и таблицы измерений. Схема «звезда» для аналитической витрины.
45. Основные конструкции SQL для получения аналитической выборки: SELECT, WHERE, ORDER BY.
46. Соединения таблиц JOIN и их применение для объединения данных предметной области.
47. Агрегатные функции, GROUP BY и HAVING при расчете аналитических показателей.
48. Подзапросы и общие табличные выражения (CTE) в аналитических запросах.
49. Оконные функции SQL: ранжирование, накопительные показатели и сравнение периодов.
50. Требования к структуре аналитической БД, обеспечивающей воспроизводимый расчет показателей и подготовку данных для моделирования.

Типовые практические задания при промежуточной аттестации

1. Для заданного описания бизнес-процесса сформулировать аналитическую задачу, определить набор необходимых данных и не менее пяти измеримых показателей.
2. По фрагменту набора данных выявить проблемы качества, выбрать способы обработки и обосновать последовательность преобразований.
3. По представленным описательным статистикам и распределениям сформулировать выводы о состоянии показателя и возможных аномалиях.
4. Рассчитать и интерпретировать заданный показатель результативности информационной системы по исходной таблице наблюдений.
5. Для прикладной ситуации сформулировать статистические гипотезы, выбрать критерий и интерпретировать предоставленное p-значение.
6. По результатам двух регрессионных моделей выбрать предпочтительную с использованием MAE/RMSE/R² и объяснить ограничения выбора.
7. По матрице ошибок рассчитать precision, recall и F1 и выбрать модель для заданной стоимости ошибок.
8. По архитектурной спецификации спроектировать структуру реляционной БД: сущности, атрибуты, ключи, связи и ограничения.
9. Нормализовать исходную таблицу до третьей нормальной формы и указать первичные и внешние ключи.
10. Написать SQL-запрос с JOIN и GROUP BY для расчета показателя по сегментам и периодам.
11. Написать SQL-запрос с оконной функцией для ранжирования или накопительного итога.
12. Спроектировать аналитическую витрину по схеме «звезда» для заданного набора показателей.
13. Разработать сквозную схему аналитического решения: источник → структура БД → подготовка → метод анализа → метрика → представление результата.