Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Гриб Владислав Валерьевич
Должность: Ректор
Дата подписания: 23.09.2026 19:10:44
Уникальный программный ключ:
637517d24e103c3db032acf37e839d98ec1c5bb2f5eb89c29abfcd7f43985447

Образовательное частное учреждение высшего образования

«МОСКОВСКИЙ УНИВЕРСИТЕТ ИМЕНИ А.С. ГРИБОЕДОВА»

(ИМПЭ им. А.С. Грибоедова)

МЕЖДУНАРОДНЫЙ ИНСТИТУТ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ И БИЗНЕС-ИНФОРМАТИКИ

УТВЕРЖДАЮ

Директор Международного института информационных технологий и бизнес-информатики

_____ Т.В. Новикова

«17» декабря 2025г.

ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

по учебной дисциплине:
ВВЕДЕНИЕ В АНАЛИЗ ДАННЫХ

по направлению подготовки/специальности:
09.03.03 Прикладная информатика
(уровень бакалавриат)

профилю/специализации:
«Анализ данных»
Формы обучения: очная, заочная

Москва

Фонд оценочных средств для дисциплины «Введение в анализ данных». Направление подготовки/специальность 09.03.03 Прикладная информатика, направленность (профиль/специализация): Анализ Данных/ Л.К. Шаймарданова. – М.: ИМПЭ им. А.С. Грибоедова.

Фонд оценочных средств является неотъемлемой частью рабочей программы дисциплины.

Разработчик: Л.К. Шаймарданова

Заведующий кафедрой: _____/ Н. Н. Загускин

1. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ДЛЯ ТЕКУЩЕГО КОНТРОЛЯ УСПЕВАЕМОСТИ, ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ОБУЧАЮЩИХСЯ ПО ДИСЦИПЛИНЕ

Настоящий Фонд оценочных средств (ФОС) по дисциплине «Введение в анализ данных» является неотъемлемым приложением к рабочей программе дисциплины (РПД) «Введение в анализ данных». На данный ФОС распространяются все реквизиты утверждения, представленные в РПД по данной дисциплине.

2. ПЕРЕЧЕНЬ ОЦЕНОЧНЫХ СРЕДСТВ

Для определения качества освоения обучающимися учебного материала по дисциплине используются следующие оценочные средства:

№ п/п	Оценочное средство	Краткая характеристика оценочного средства	Представление оценочного средства в ФОС
1	Тестирование	Вид контроля, позволяющий оценить изученный теоретический материал.	Вопросы для проведения тестирования
2	Практические задания	Вид контроля, позволяющий оценить умение обучающегося применять осваиваемую компетенцию в практических ситуациях и при решении производственных задач	Задания к практическому занятию
3	Контрольная работа	Вид контроля, позволяющий определить результат освоения компетенций по дисциплине в рамках рассматриваемой темы, оцениваемый с помощью соответствующих индикаторов достижения компетенций	Задания контрольной работы
4	Самостоятельная работа	Вид контроля, позволяющий оценить проработку теоретического материала, изучение рекомендуемой литературы, выполнение практико-ориентированных заданий (заполнение таблиц, проведение сравнительного анализа, составление схем и др.), решение практических задач, создание презентаций, написание рефератов, подборку нормативного и иного материала и выполнение других заданий	Задания самостоятельной работы
5	Курсовая работа	Вид контроля, позволяющий выявить степень владения базовыми знаниями, умениями и навыками, необходимыми для обучения, и определить уровень владения новым материалом	Индивидуальные задания (темы) для курсовой работы
6	Зачет/Зачет с оценкой/Экзамен	Вид контроля, позволяющий выявить степень овладения знаниями, умениями и навыками, необходимыми для дальнейшего освоения образовательной программы подготовки	Вопросы для подготовки к зачету/зачету с оценкой/экзамену

3. ПАСПОРТ ФОНДА ОЦЕНОЧНЫХ СРЕДСТВ ДИСЦИПЛИНЕ

Разработчик	Л.К. Шаймарданова
Кафедра	Прикладной информатики и информационных технологий
Наименование дисциплины	Введение в анализ данных

Факультет / институт	Институт международной экономики, лидерства и менеджмента
Направление подготовки / специальность	09.03.03 Прикладная информатика
Количество вопросов в оценочных заданиях (диапазон)	От 10 до 20
Общее время тестирования (мин)	90
Общее количество вопросов/заданий в ФОС	20
Размещенность на сайте Университета примерного перечня вопросов, заданий ФОС – для подготовки обучающихся к прохождению оценки (да / нет)	Да

3.1. Модели контролируемых компетенций:

Код компетенции	Формулировка компетенции	Индикаторы достижения компетенции
ПК-5	Способен осуществлять проектирование структур данных	**ИПК-5.1 Знать:** Методы и средства проектирования компьютерного программного обеспечения **ИПК-5.2 Уметь:** Применять методы и средства проектирования компьютерного программного обеспечения, структур данных, баз данных, программных интерфейсов
ПК-6	Способен применять методы и средства проектирования компьютерного программного обеспечения	**ИПК-6.1 Знать:** Методы и средства проектирования баз данных **ИПК-6.2 Уметь:** Применять методы и средства проектирования компьютерного программного обеспечения, структур данных, баз данных, программных интерфейсов

4. СОДЕРЖАНИЕ ОЦЕНОЧНЫХ СРЕДСТВ ТЕКУЩЕГО КОНТРОЛЯ

4.1. ТИПОВЫЕ ТЕСТОВЫЕ ЗАДАНИЯ.

Тесты содержат набор вопросов, в полном объеме охватывающие изученный теоретический материал по указанной теме. Выполнение тестов позволяет определить результат освоения компетенций по дисциплине в рамках рассматриваемой темы, оцениваемый с помощью соответствующих индикаторов достижения компетенций. Индивидуальный тестовый сеанс для каждого обучающегося формируется по специальному алгоритму, обеспечивающему заданную тематическую структуру и пропорциональное наличие вопросов разного типа и сложности.

При формировании тестов необходимо использовать задания следующих типов:

Тип задания 1. Задания закрытого типа на установление соответствия.

Тип задания 2. Задания закрытого типа на установление последовательности.

Тип задания 3. Задания комбинированного типа, предполагающие выбор одного правильного ответа из предложенных

Тип задания 4. Задания комбинированного типа, предполагающие выбор нескольких ответов из предложенных

Тип задания 5. Задания открытого типа с развернутым ответом.

Тема	Код и формулировка компетенции	Индикаторы достижения компетенции	Тест по теме
Тема 1.1 Процесс анализа данных.	ПК- 5 Способен осуществлять проектирование структур данных баз данных ПК-6 Способен применять методы и средства проектирования компьютерного программного обеспечения	**ИПК-5.1 Знать:** Методы и средства проектирования компьютерного программного обеспечения **ИПК-6.1 Знать:** Методы и средства проектирования баз данных	**Задание 1.1** Какие утверждения корректно отражают сущность процесса анализа данных? Варианты ответа (выберите все верные): 1. Анализ данных направлен на извлечение знаний и закономерностей из данных 2. Анализ данных заключается только в хранении данных в базах данных 3. Анализ данных включает этапы сбора, очистки, обработки и интерпретации данных 4. Анализ данных не предполагает формулировки цели исследования 5. Анализ данных используется для поддержки принятия управленческих и прикладных решений Ответ: 1, 3, 5 Пояснение: 1 — верно, основной целью анализа данных является получение знаний и выводов. 3 — верно, анализ данных представляет собой многоэтапный процесс. 5 — верно, результаты анализа используются для принятия решений. **Задание 1.2** Какие утверждения корректно характеризуют уровни информации и стандарты анализа данных? Варианты ответа (выберите все верные): 1. Данные представляют собой необработанные факты и наблюдения 2. Информация формируется в результате интерпретации данных в контексте

Тема	Код и формулировка компетенции	Индикаторы достижения компетенции	Тест по теме
			3. Знания находятся на нижнем уровне информационной пирамиды 4. Стандарты анализа данных обеспечивают воспроизводимость и единый подход к анализу 5. CRISP-DM является примером стандарта процесса анализа данных Ответ: 1, 2, 4, 5 Пояснение: 1 — верно, данные не несут смысла без обработки. 2 — верно, информация — это осмысленные данные. 4 — верно, стандарты регламентируют процесс анализа. 5 — верно, CRISP-DM — общепринятый стандарт жизненного цикла анализа данных. **Задание 1.3** Расположите этапы анализа данных в логически корректной последовательности. Этапы: 1. Моделирование 2. Понимание задачи (целей анализа) 3. Подготовка данных 4. Понимание и первичный анализ данных 5. Оценка результатов 6. Внедрение результатов Ответ: 243156 Пояснение: Сначала формулируется цель анализа, затем изучаются данные, после чего они подготавливаются. Далее строятся модели, оценивается их качество и только после этого результаты внедряются.

Тема	Код и формулировка компетенции	Индикаторы достижения компетенции	Тест по теме
			Задание 1.4 Расположите уровни информационной пирамиды от базового к наиболее высокому. Уровни пирамиды: 1. Знания 2. Данные 3. Мудрость 4. Информация Ответ: 2413 Пояснение: Данные являются основой, из них формируется информация, затем — знания, а высший уровень — мудрость как способность принимать обоснованные решения. **Задание 1.5** Установите соответствие между уровнем информационной пирамиды и его описанием. Список 1: 1. Данные 2. Информация 3. Знания 4. Мудрость Список 2: A. Осмысленные выводы и рекомендации для принятия решений B. Необработанные факты и значения C. Выявленные закономерности и обобщения D. Данные, интерпретированные в определённом контексте

Тема	Код и формулировка компетенции	Индикаторы достижения компетенции	Тест по теме
			Ответ: 1B, 2D, 3C, 4A **Задание 1.6** Установите соответствие между стандартом и его характеристикой. Список 1: 1. CRISP-DM 2. KDD 3. SEMMA Список 2: A. Стандарт, ориентированный на этапы выборки, моделирования и оценки B. Процесс извлечения знаний из баз данных C. Универсальная модель жизненного цикла анализа данных Ответ: 1C, 2B, 3A **Задание 1.7** Какова основная цель анализа данных? Варианты ответа: 1. Хранение больших объёмов информации 2. Извлечение знаний и поддержка принятия решений 3. Создание информационных систем 4. Замена экспертных оценок без проверки Ответ: 2 Пояснение: Анализ данных направлен на получение знаний и обоснование решений, а не на хранение или разработку систем.

Тема	Код и формулировка компетенции	Индикаторы достижения компетенции	Тест по теме
			Задание 1.8 Какой стандарт наиболее широко используется для описания жизненного цикла анализа данных? Варианты ответа: 1. UML 2. SQL 3. CRISP-DM 4. HTTP Ответ: 3 Пояснение: CRISP-DM — общепринятый стандарт, описывающий этапы анализа данных от постановки задачи до внедрения. **Задание 1.9** Как называется уровень информационной пирамиды, на котором данные приобретают смысл в контексте? Ответ: Информация Пояснение: Информация формируется в результате интерпретации данных и добавления контекста. **Задание 1.10** Как называется стандарт процесса анализа данных, включающий этапы понимания задачи, понимания данных, подготовки данных, моделирования, оценки и внедрения? Ответ: CRISP-DM

Тема	Код и формулировка компетенции	Индикаторы достижения компетенции	Тест по теме
			Пояснение: CRISP-DM описывает полный жизненный цикл анализа данных и широко применяется в академической и прикладной практике.
Тема 2.1 Предобработка и подготовка данных. Обработка данных с помощью пакета Pandas	ПК- 5 Способен осуществлять проектирование структур данных баз данных ПК-6 Способен применять методы и средства проектирования компьютерного программного обеспечения	**ИПК-5.1 Знать:** Методы и средства проектирования компьютерного программного обеспечения **ИПК-5.2 Уметь:** Применять методы и средства проектирования компьютерного программного обеспечения, структур данных, баз данных, программных интерфейсов **ИПК-6.1 Знать:** Методы и средства проектирования баз данных **ИПК-6.2 Уметь:** Применять методы и средства проектирования компьютерного программного обеспечения, структур данных, баз данных, программных интерфейсов	**Задание 2.1** Какие утверждения корректно описывают структуры данных Series и DataFrame библиотеки Pandas? Варианты ответа (выберите все верные): 1. Series представляет собой одномерную структуру данных с индексами 2. DataFrame является двумерной табличной структурой данных 3. Series может содержать данные разных типов в одном объекте 4. DataFrame состоит из набора объектов типа Series 5. Series не поддерживает работу с индексами Ответ: 1, 2, 4 Пояснение: 1 — верно, Series — одномерный массив с индексами. 2 — верно, DataFrame представляет собой таблицу (строки и столбцы). 4 — верно, каждый столбец DataFrame является объектом Series. **Задание 2.2** Какие методы Pandas используются для предварительного анализа структуры данных и визуализации? Варианты ответа (выберите все верные): 1. head() 2. info() 3. describe() 4. plot() 5. merge() Ответ: 1, 2, 3, 4

Тема	Код и формулировка компетенции	Индикаторы достижения компетенции	Тест по теме
			Пояснение: 1 — вывод первых строк набора данных. 2 — отображает информацию о типах данных и пропусках. 3 — рассчитывает описательные статистики. 4 — используется для визуализации данных. **Задание 2.3** Расположите этапы работы с данными в Pandas в логически корректной последовательности. Этапы: 1. Визуальный анализ данных 2. Загрузка данных в DataFrame 3. Анализ структуры данных (info(), head()) 4. Очистка и преобразование данных Ответ: 2341 Пояснение: Сначала данные загружаются, затем изучается их структура, после чего выполняется очистка, и только затем проводится визуальный анализ. **Задание 2.4** Упорядочите действия при первичном анализе нового набора данных в Pandas. Действия: 1. Построение графиков 2. Просмотр первых строк данных 3. Получение информации о типах данных и пропусках 4. Расчёт описательных статистик

Тема	Код и формулировка компетенции	Индикаторы достижения компетенции	Тест по теме
			Ответ: 2341 Пояснение: Сначала изучается содержимое таблицы, затем её структура, далее числовые характеристики, и только после этого — визуализация. **Задание 2.5** Установите соответствие между методом Pandas и его назначением. Список 1: 1. head() 2. info() 3. describe() 4. plot() Список 2: A. Визуализация данных B. Вывод общей информации о структуре DataFrame C. Отображение первых строк таблицы D. Расчёт описательных статистик Ответ: 1C, 2B, 3D, 4A **Задание 2.6** Установите соответствие между структурой данных Pandas и её характеристикой. Список 1: 1. Series 2. DataFrame Список 2: A. Двумерная табличная структура данных

Тема	Код и формулировка компетенции	Индикаторы достижения компетенции	Тест по теме
			B. Одномерный массив с индексами Ответ: 1B, 2A **Задание 2.7** Какой метод Pandas позволяет получить информацию о типах данных столбцов и количестве пропусков? Варианты ответа: 1. head() 2. describe() 3. info() 4. plot() Ответ: 3 Пояснение: Метод info() выводит типы данных, количество ненулевых значений и объём памяти. **Задание 2.8** Какой метод Pandas используется для построения графиков на основе данных DataFrame? Варианты ответа: 1. visualize() 2. draw() 3. plot() 4. show() Ответ: 3 Пояснение:

Тема	Код и формулировка компетенции	Индикаторы достижения компетенции	Тест по теме
			Метод plot() предоставляет встроенные средства визуализации на основе Matplotlib. **Задание 2.9** Как называется двумерная табличная структура данных библиотеки Pandas? Ответ: DataFrame Пояснение: DataFrame предназначен для хранения и обработки табличных данных с индексами и столбцами. **Задание 2.10** Какой метод Pandas используется для вычисления основных описательных статистик числовых столбцов? Ответ: describe() Пояснение: Метод describe() рассчитывает среднее, медиану, стандартное отклонение и другие статистические показатели.

Критерии оценивания тестового задания:

Оценка	Критерии оценивания
отлично	от 90 до 100 % правильно выполненных заданий
хорошо	от 70 до 89 % правильно выполненных заданий
удовлетворительно	от 50 до 69 % правильно выполненных заданий
неудовлетворительно	менее 50 % правильно выполненных заданий

4.2 ТИПОВЫЕ ПРАКТИЧЕСКИЕ ЗАДАНИЯ

Практические задания должны отражать умение обучающегося применять осваиваемую компетенцию в практических ситуациях и при решении производственных задач.

Тема	Код и формулировка компетенции	Индикаторы достижения компетенции	Задания к практическому занятию

Тема 1.1 Процесс анализа данных.	ПК- 5 Способен осуществлять проектирование структур данных баз данных
	
	ПК-6 Способен применять методы и средства проектирования компьютерного программного обеспечения
		ИПК-5.1 Знать:
		
		Методы и средства проектирования компьютерного программного обеспечения
		
		ИПК-5.2 Уметь:
		
		Применять методы и средства проектирования компьютерного программного обеспечения, структур данных, баз данных, программных интерфейсов
		
		ИПК-6.1 Знать:
		
		Методы и средства проектирования баз данных
		
		ИПК-6.2 Уметь:
		
		Применять методы и средства проектирования компьютерного программного обеспечения, структур данных, баз данных, программных интерфейсов
			Задание 1.1
			Создайте DataFrame с пропущенными значениями и дубликатами. Выполните заполнение пропусков средним значением и удалите дубликаты.
			Дан набор данных:
			● ages = [25, 30, None, 30]
			● salaries = [50000, 60000, 55000, 60000]
			Задачи:
			1. Создайте DataFrame из этих списков.
			2. Заполните пропущенное значение в столбце age средним по столбцу.
			3. Удалите дублирующиеся строки.
			4. Выведите итоговый DataFrame.
			
			Ответ
			`import pandas as pd`
			`import numpy as np`
			
			`df = pd.DataFrame({`
			`  'age': [25, 30, np.nan, 30],`
			`  'salary': [50000, 60000, 55000, 60000]`
			`})`
			
			`df['age'] = df['age'].fillna(df['age'].mean())`
			`df = df.drop_duplicates()`
			
			`print(df)`
			
			Задание 1.2
			Обнаружьте и удалите аномальные значения в столбце salary, превышающие 100000.

Тема	Код и формулировка компетенции	Индикаторы достижения компетенции	Задания к практическому занятию

			Дан список заработных плат:
			● salaries = [40000, 50000, 120000, 60000]
			Задачи:
			1. Создайте DataFrame.
			2. Считайте аномальными все значения > 100000 и удалите их.
			3. Выведите очищенный DataFrame.
			
			Ответ:
			`import pandas as pd`
			
			`df = pd.DataFrame({'salary': [40000, 50000, 120000, 60000]})`
			`df = df[df['salary'] <= 100000]`
			
			`print(df)`
			
			Задание 2.1
			Создайте объект Series, добавьте новый элемент и удалите элемент с индексом 1.
			Дан список чисел:
			● numbers = [10, 20, 30]
			Задачи:
			1. Создайте Series из этого списка.
			2. Добавьте элемент со значением 40 и индексом 3.
			3. Удалите элемент с индексом 1.
			4. Выведите итоговый Series.
			
			Ответ:
			`import pandas as pd`
			
			`s = pd.Series([10, 20, 30])`
			`s.loc[3] = 40`
			`s = s.drop(1)`

Тема	Код и формулировка компетенции	Индикаторы достижения компетенции	Задания к практическому занятию
			print(s) **Задание 2.2** Создайте DataFrame, добавьте вычисляемый столбец и удалите указанный столбец. Даны списки: ● A = [1, 2, 3] ● B = [4, 5, 6] Задачи: 1. Создайте DataFrame из этих списков. 2. Добавьте столбец C, равный сумме A и B. 3. Удалите столбец B. 4. Выведите итоговый DataFrame. Ответ: import pandas as pd df = pd.DataFrame({ 'A': [1, 2, 3], 'B': [4, 5, 6] }) df['C'] = df['A'] + df['B'] df = df.drop(columns=['B']) print(df) **Задание 3.1** Выполните срез данных, выбрав строки, где значение score больше 80. Даны списки: ● names = ['A', 'B', 'C']

Тема	Код и формулировка компетенции	Индикаторы достижения компетенции	Задания к практическому занятию

			● scores = [75, 85, 90]
			Задачи:
			1. Создайте DataFrame.
			2. Выберите строки, где score > 80.
			3. Выведите результат.
			
			Ответ:
			`import pandas as pd`
			
			`df = pd.DataFrame({`
			`  'name': ['A', 'B', 'C'],`
			`  'score': [75, 85, 90]`
			`})`
			
			`print(df[df['score'] > 80])`
			
			Задание 3.2
			Выполните группировку по столбцу group и проанализируйте структуру данных.
			Даны списки:
			● groups = ['X', 'X', 'Y']
			● values = [10, 20, 30]
			Задачи:
			1. Создайте DataFrame.
			2. Сгруппируйте данные по group и вычислите среднее значение value.
			3. Выведите информацию о структуре DataFrame.
			
			Ответ:
			`import pandas as pd`
			
			`df = pd.DataFrame({`
			`  'group': ['X', 'X', 'Y'],`

Тема	Код и формулировка компетенции	Индикаторы достижения компетенции	Задания к практическому занятию
			'value': [10, 20, 30] }) print(df.groupby('group').mean()) print(df.info()) **Задание 4.1** Рассчитайте основные статистические характеристики числового столбца. Дан список: ● data = [10, 20, 30, 40] Задачи: 1. Создайте DataFrame. 2. Вычислите основные статистические характеристики (mean, median, std) столбца data. Ответ: import pandas as pd df = pd.DataFrame({'data': [10, 20, 30, 40]}) print(df['data'].describe()) **Задание 4.2** Проведите корреляционный анализ двух числовых признаков. Даны списки: ● x = [1, 2, 3, 4] ● y = [2, 3, 5, 7] Задачи: 1. Создайте DataFrame. 2. Вычислите корреляцию между x и y. Ответ: import pandas as pd

Тема	Код и формулировка компетенции	Индикаторы достижения компетенции	Задания к практическому занятию
			df = pd.DataFrame({ 'x': [1, 2, 3, 4], 'y': [2, 3, 5, 7] }) print(df.corr())
Тема 2.1 Предобработка и подготовка данных. Обработка данных с помощью пакета Pandas	ПК- 5 Способен осуществлять проектирование структур данных баз данных ПК-6 Способен применять методы и средства проектирования компьютерного программного обеспечения	**ИПК-5.1 Знать:** Методы и средства проектирования компьютерного программного обеспечения **ИПК-5.2 Уметь:** Применять методы и средства проектирования компьютерного программного обеспечения, структур данных, баз данных, программных интерфейсов **ИПК-6.1 Знать:** Методы и средства проектирования баз данных **ИПК-6.2 Уметь:** Применять методы и средства проектирования компьютерного программного обеспечения, структур данных, баз данных, программных интерфейсов	**Задание 5.1** Постройте гистограмму и линейный график для заданных данных. Дан список: ● values = [10, 20, 30, 40, 50] Задачи: 1. Создайте DataFrame. 2. Постройте гистограмму и линейный график столбца values. Ответ: import pandas as pd import matplotlib.pyplot as plt df = pd.DataFrame({'value': [10, 20, 30, 40, 50]}) df['value'].plot(kind='hist') plt.show() df['value'].plot(kind='line') plt.show() **Задание 5.2** Постройте круговую и точечную диаграммы. Даны списки: ● categories = ['A', 'B', 'C']

Тема	Код и формулировка компетенции	Индикаторы достижения компетенции	Задания к практическому занятию

			● counts = [10, 20, 30]
			Задачи:
			1. Создайте DataFrame.
			2. Постройте круговую и точечную диаграммы.
			
			Ответ:
			`import pandas as pd`
			`import matplotlib.pyplot as plt`
			
			`df = pd.DataFrame({`
			`'category': ['A', 'B', 'C'],`
			`'count': [10, 20, 30]`
			`})`
			
			`df.set_index('category').plot(kind='pie', y='count')`
			`plt.show()`
			
			`df.plot(kind='scatter', x='count', y='count')`
			`plt.show()`
			
			Задание 6.1
			Выполните описательный статистический анализ данных.
			Даны списки:
			● A = [5, 10, 15]
			● B = [3, 6, 9]
			Задачи:
			1. Создайте DataFrame.
			2. Выполните описательный статистический анализ.
			
			Ответ:
			`import pandas as pd`
			
			`df = pd.DataFrame({`
			`'A': [5, 10, 15],`
			`'B': [3, 6, 9]`

Тема	Код и формулировка компетенции	Индикаторы достижения компетенции	Задания к практическому занятию
			}) print(df.describe()) **Задание 6.2** Выполните графический анализ двух числовых признаков. Даны списки: ● x = [1, 2, 3, 4] ● y = [4, 3, 2, 1] Задачи: 1. Создайте DataFrame. 2. Постройте диаграмму рассеяния для x и y. Ответ: import pandas as pd import matplotlib.pyplot as plt df = pd.DataFrame({ 'x': [1, 2, 3, 4], 'y': [4, 3, 2, 1] }) df.plot(kind='scatter', x='x', y='y') plt.show() **Задание 7.1** Удалите пропуски и выполните статистический анализ. Дан список: ● values = [10, None, 30, 40] Задачи: 1. Создайте DataFrame. 2. Удалите пропуски. 3. Выведите описательную статистику.

Тема	Код и формулировка компетенции	Индикаторы достижения компетенции	Задания к практическому занятию
			Ответ: import pandas as pd import numpy as np df = pd.DataFrame({'value': [10, np.nan, 30, 40]}) df = df.dropna() print(df.describe()) **Задание 7.2** Постройте график распределения очищенных данных. Дан список: ● values = [10, 20, 30, 40] Задачи: 1. Постройте гистограмму распределения Ответ: import pandas as pd import matplotlib.pyplot as plt df = pd.DataFrame({'value': [10, 20, 30, 40]}) df['value'].plot(kind='hist') plt.show() **Задание 8.1** Сформируйте итоговую таблицу со средними значениями по группам. Даны списки: ● groups = ['A', 'A', 'B', 'B'] ● values = [10, 20, 30, 40] Задачи: 1. Создайте DataFrame.

Тема	Код и формулировка компетенции	Индикаторы достижения компетенции	Задания к практическому занятию

			2. Сгруппируйте данные по group и вычислите среднее значение value.
			
			Ответ:
			`import pandas as pd`
			
			`df = pd.DataFrame({`
			`  'group': ['A', 'A', 'B', 'B'],`
			`  'value': [10, 20, 30, 40]`
			`})`
			
			`result = df.groupby('group').mean()`
			`print(result)`
			
			Задание 8.2
			Визуализируйте итоговые результаты анализа.
			● Даны списки:
			group: ['A', 'B'],
			● value: [15, 35]
			Задача:
			1. Постройте столбчатую диаграмму средних значений по группам.
			
			Ответ:
			`import pandas as pd`
			`import matplotlib.pyplot as plt`
			
			`df = pd.DataFrame({`
			`  'group': ['A', 'B'],`
			`  'value': [15, 35]`
			`})`
			
			`df.set_index('group').plot(kind='bar')`
			`plt.show()`

Критерии оценивания практических занятий:

Оценка	Критерии оценивания
отлично	Выставляется, если обучающийся умеет увязывать теорию с практикой (решает задачи, формулирует выводы, умеет пояснить полученные результаты), владеет понятийным аппаратом, полно и глубоко овладел материалом по заданной теме, обосновывает свои суждения и даёт правильные ответы на вопросы преподавателя
хорошо	Выставляется, если обучающийся умеет увязывать теорию с практикой (решает задачи и формулирует выводы, умеет пояснить полученные результаты), владеет понятийным аппаратом, полно и глубоко овладел материалом по заданной теме, но содержание ответов имеют некоторые неточности и требуют уточнения и комментария со стороны преподавателя
удовлетворительно	Выставляется, если обучающийся знает и понимает материал по заданной теме, но изложение неполное, непоследовательное, допускаются неточности в определении понятий, студент не может обосновать свои ответы на уточняющие вопросы преподавателя
неудовлетворительно	Выставляется, если обучающийся допускает ошибки в определении понятий, искажающие их смысл, беспорядочно и неуверенно излагает материал. Делает ошибки в ответах на уточняющие вопросы преподавателя

4.3. ТИПОВЫЕ ЗАДАНИЯ ДЛЯ КОНТРОЛЬНЫХ РАБОТ

Контрольные работы содержат несколько практических заданий по индивидуальным вариантам, в полном объеме охватывающих изученный материал по указанной теме. Выполнение контрольных работ позволяет определить результат освоения компетенций по дисциплине в рамках рассматриваемой темы, оцениваемый с помощью соответствующих индикаторов достижения компетенций.

Тема	Код и формулировка компетенции	Индикаторы достижения компетенции	Типовой вариант контрольной работы
Тема 1.1 Процесс анализа данных.	ПК- 5 Способен осуществлять проектирование структур данных баз данных ПК-6 Способен применять методы и средства проектировани	**ИПК-5.1 Знать:** Методы и средства проектирования компьютерного программного обеспечения **ИПК-5.2 Уметь:** Применять методы и средства проектирования компьютерного программного обеспечения, структур данных, баз данных, программных интерфейсов **ИПК-6.1 Знать:** Методы и средства проектирования баз данных	**Задание 1.1** Какие утверждения корректно отражают сущность процесса анализа данных? Варианты ответа (выберите все верные): 1. Анализ данных направлен на извлечение знаний и закономерностей из данных 2. Анализ данных заключается только в хранении данных в базах данных 3. Анализ данных включает этапы сбора, очистки, обработки и интерпретации данных 4. Анализ данных не предполагает формулировки цели исследования 5. Анализ данных используется для поддержки принятия управленческих и прикладных решений

Тема	Код и формулировка компетенции	Индикаторы достижения компетенции	Типовой вариант контрольной работы
	я компьютерного программного обеспечения	**ИПК-6.2 Уметь:** Применять методы и средства проектирования компьютерного программного обеспечения, структур данных, баз данных, программных интерфейсов	Ответ: 1, 3, 5 Пояснение: 1 — верно, основной целью анализа данных является получение знаний и выводов. 3 — верно, анализ данных представляет собой многоэтапный процесс. 5 — верно, результаты анализа используются для принятия решений. **Задание 1.2** Какие утверждения корректно характеризуют уровни информации и стандарты анализа данных? Варианты ответа (выберите все верные): 1. Данные представляют собой необработанные факты и наблюдения 2. Информация формируется в результате интерпретации данных в контексте 3. Знания находятся на нижнем уровне информационной пирамиды 4. Стандарты анализа данных обеспечивают воспроизводимость и единый подход к анализу 5. CRISP-DM является примером стандарта процесса анализа данных Ответ: 1, 2, 4, 5 Пояснение: 1 — верно, данные не несут смысла без обработки. 2 — верно, информация — это осмысленные данные. 4 — верно, стандарты регламентируют процесс анализа. 5 — верно, CRISP-DM — общепринятый стандарт жизненного цикла анализа данных. **Задание 1.3** Расположите этапы анализа данных в логически корректной

Тема	Код и формулировка компетенции	Индикаторы достижения компетенции	Типовой вариант контрольной работы
			последовательности. Этапы: 1. Моделирование 2. Понимание задачи (целей анализа) 3. Подготовка данных 4. Понимание и первичный анализ данных 5. Оценка результатов 6. Внедрение результатов Ответ: 243156 Пояснение: Сначала формулируется цель анализа, затем изучаются данные, после чего они подготавливаются. Далее строятся модели, оценивается их качество и только после этого результаты внедряются. **Задание 1.4** Расположите уровни информационной пирамиды от базового к наиболее высокому. Уровни пирамиды: 1. Знания 2. Данные 3. Мудрость 4. Информация Ответ: 2413 Пояснение: Данные являются основой, из них формируется информация, затем — знания, а высший уровень — мудрость как способность принимать обоснованные решения.

Тема	Код и формулировка компетенции	Индикаторы достижения компетенции	Типовой вариант контрольной работы
			Задание 1.5 Установите соответствие между уровнем информационной пирамиды и его описанием. Список 1: 1. Данные 2. Информация 3. Знания 4. Мудрость Список 2: A. Осмысленные выводы и рекомендации для принятия решений B. Необработанные факты и значения C. Выявленные закономерности и обобщения D. Данные, интерпретированные в определённом контексте Ответ: 1B, 2D, 3C, 4A **Задание 1.6** Установите соответствие между стандартом и его характеристикой. Список 1: 1. CRISP-DM 2. KDD 3. SEMMA Список 2: A. Стандарт, ориентированный на этапы выборки, моделирования и оценки B. Процесс извлечения знаний из баз данных C. Универсальная модель жизненного цикла анализа данных Ответ: 1C, 2B, 3A

Тема	Код и формулировка компетенции	Индикаторы достижения компетенции	Типовой вариант контрольной работы
			Задание 1.7 Какова основная цель анализа данных? Варианты ответа: 1. Хранение больших объёмов информации 2. Извлечение знаний и поддержка принятия решений 3. Создание информационных систем 4. Замена экспертных оценок без проверки Ответ: 2 Пояснение: Анализ данных направлен на получение знаний и обоснование решений, а не на хранение или разработку систем. **Задание 1.8** Какой стандарт наиболее широко используется для описания жизненного цикла анализа данных? Варианты ответа: 1. UML 2. SQL 3. CRISP-DM 4. HTTP Ответ: 3 Пояснение: CRISP-DM — общепринятый стандарт, описывающий этапы анализа данных от постановки задачи до внедрения. **Задание 1.9**

Тема	Код и формулировка компетенции	Индикаторы достижения компетенции	Типовой вариант контрольной работы
			Как называется уровень информационной пирамиды, на котором данные приобретают смысл в контексте? Ответ: Информация Пояснение: Информация формируется в результате интерпретации данных и добавления контекста. **Задание 1.10** Как называется стандарт процесса анализа данных, включающий этапы понимания задачи, понимания данных, подготовки данных, моделирования, оценки и внедрения? Ответ: CRISP-DM Пояснение: CRISP-DM описывает полный жизненный цикл анализа данных и широко применяется в академической и прикладной практике.
Тема 2.1 Предобработка и подготовка данных. Обработка данных с помощью пакета Pandas	ПК- 5 Способен осуществлять проектирование структур данных баз данных ПК-6 Способен применять методы и средства проектирования	**ИПК-5.1 Знать:** Методы и средства проектирования компьютерного программного обеспечения **ИПК-5.2 Уметь:** Применять методы и средства проектирования компьютерного программного обеспечения, структур данных, баз данных, программных интерфейсов **ИПК-6.1 Знать:** Методы и средства проектирования баз данных **ИПК-6.2 Уметь:** Применять методы и средства проектирования компьютерного	**Задание 2.1** Какие утверждения корректно описывают структуры данных Series и DataFrame библиотеки Pandas? Варианты ответа (выберите все верные): 1. Series представляет собой одномерную структуру данных с индексами 2. DataFrame является двумерной табличной структурой данных 3. Series может содержать данные разных типов в одном объекте 4. DataFrame состоит из набора объектов типа Series 5. Series не поддерживает работу с индексами Ответ: 1, 2, 4 Пояснение:

Тема	Код и формулировка компетенции	Индикаторы достижения компетенции	Типовой вариант контрольной работы
	компьютерного программного обеспечения	программного обеспечения, структур данных, баз данных, программных интерфейсов	1 — верно, Series — одномерный массив с индексами. 2 — верно, DataFrame представляет собой таблицу (строки и столбцы). 4 — верно, каждый столбец DataFrame является объектом Series. **Задание 2.2** Какие методы Pandas используются для предварительного анализа структуры данных и визуализации? Варианты ответа (выберите все верные): 1. head() 2. info() 3. describe() 4. plot() 5. merge() Ответ: 1, 2, 3, 4 Пояснение: 1 — вывод первых строк набора данных. 2 — отображает информацию о типах данных и пропусках. 3 — рассчитывает описательные статистики. 4 — используется для визуализации данных. **Задание 2.3** Расположите этапы работы с данными в Pandas в логически корректной последовательности. Этапы: 1. Визуальный анализ данных 2. Загрузка данных в DataFrame 3. Анализ структуры данных (info(), head()) 4. Очистка и преобразование данных

Тема	Код и формулировка компетенции	Индикаторы достижения компетенции	Типовой вариант контрольной работы
			Ответ: 2341 Пояснение: Сначала данные загружаются, затем изучается их структура, после чего выполняется очистка, и только затем проводится визуальный анализ. **Задание 2.4** Упорядочите действия при первичном анализе нового набора данных в Pandas. Действия: 1. Построение графиков 2. Просмотр первых строк данных 3. Получение информации о типах данных и пропусках 4. Расчёт описательных статистик Ответ: 2341 Пояснение: Сначала изучается содержимое таблицы, затем её структура, далее числовые характеристики, и только после этого — визуализация. **Задание 2.5** Установите соответствие между методом Pandas и его назначением. Список 1: 1. head() 2. info() 3. describe() 4. plot() Список 2: A. Визуализация данных

Тема	Код и формулировка компетенции	Индикаторы достижения компетенции	Типовой вариант контрольной работы
			B. Вывод общей информации о структуре DataFrame C. Отображение первых строк таблицы D. Расчёт описательных статистик Ответ: 1C, 2B, 3D, 4A **Задание 2.6** Установите соответствие между структурой данных Pandas и её характеристикой. Список 1: 1. Series 2. DataFrame Список 2: A. Двумерная табличная структура данных B. Одномерный массив с индексами Ответ: 1B, 2A **Задание 2.7** Какой метод Pandas позволяет получить информацию о типах данных столбцов и количестве пропусков? Варианты ответа: 1. head() 2. describe() 3. info() 4. plot() Ответ: 3 Пояснение:

Тема	Код и формулировка компетенции	Индикаторы достижения компетенции	Типовой вариант контрольной работы
			Метод info() выводит типы данных, количество ненулевых значений и объём памяти. **Задание 2.8** Какой метод Pandas используется для построения графиков на основе данных DataFrame? Варианты ответа: 1. visualize() 2. draw() 3. plot() 4. show() Ответ: 3 Пояснение: Метод plot() предоставляет встроенные средства визуализации на основе Matplotlib. **Задание 2.9** Как называется двумерная табличная структура данных библиотеки Pandas? Ответ: DataFrame Пояснение: DataFrame предназначен для хранения и обработки табличных данных с индексами и столбцами. **Задание 2.10** Какой метод Pandas используется для вычисления основных описательных статистик числовых столбцов? Ответ: describe()

Тема	Код и формулировка компетенции	Индикаторы достижения компетенции	Типовой вариант контрольной работы
			Пояснение: Метод describe() рассчитывает среднее, медиану, стандартное отклонение и другие статистические показатели.

Критерии оценивания контрольной работы:

Оценка	Критерии оценивания
отлично	Выставляется, если обучающийся умеет увязывать теорию с практикой (решает задачи, формулирует выводы, умеет пояснить полученные результаты), владеет понятийным аппаратом, полно и глубоко овладел материалом по заданной теме, обосновывает свои суждения и даёт правильные ответы на вопросы преподавателя
хорошо	Выставляется, если обучающийся умеет увязывать теорию с практикой (решает задачи и формулирует выводы, умеет пояснить полученные результаты), владеет понятийным аппаратом, полно и глубоко овладел материалом по заданной теме, но содержание ответов имеют некоторые неточности и требуют уточнения и комментария со стороны преподавателя.
удовлетворительно	Выставляется, если обучающийся знает и понимает материал по заданной теме, но изложение неполное, непоследовательное, допускаются неточности в определении понятий, студент не может обосновать свои ответы на уточняющие вопросы преподавателя
неудовлетворительно	Выставляется, если обучающийся допускает ошибки в определении понятий, искажающие их смысл, беспорядочно и неуверенно излагает материал. Делает ошибки в ответах на уточняющие вопросы преподавателя

4.4. ТИПОВЫЕ ЗАДАНИЯ ДЛЯ САМОСТОЯТЕЛЬНОЙ РАБОТЫ

Самостоятельная работа включает в себя проработку теоретического материала, изучение рекомендуемой литературы, выполнение практико-ориентированных заданий (заполнение таблиц, проведение сравнительного анализа, составление схем и др.), решение практических задач, создание презентаций, написание рефератов, подборка нормативного и иного материала и выполнение других заданий.

Тема	Код и формулировка компетенции	Индикаторы достижения компетенции	Типовое задание для самостоятельной работы
Тема 1.1	ПК- 5 Способен	**ИПК-5.1 Знать:** Методы и средства	Самостоятельно изучите тему «Стандарты в области анализа

Тема	Код и формулировка компетенции	Индикаторы достижения компетенции	Типовое задание для самостоятельной работы
Процесс анализа данных.	осуществлять проектирование структур данных баз данных ПК-6 Способен применять методы и средства проектирования компьютерного программного обеспечения	проектирования компьютерного программного обеспечения **ИПК-5.2 Уметь:** Применять методы и средства проектирования компьютерного программного обеспечения, структур данных, баз данных, программных интерфейсов **ИПК-6.1 Знать:** Методы и средства проектирования баз данных **ИПК-6.2 Уметь:** Применять методы и средства проектирования компьютерного программного обеспечения, структур данных, баз данных, программных интерфейсов	данных» и подготовьте краткий доклад по изученному материалу Что следует отразить: В реферате на тему можно рассмотреть различные аспекты, связанные с стандартизацией в этой области. Технические стандарты для сбора, хранения, обработки и анализа больших данных. Они устанавливают общие принципы и правила для разработки и внедрения технических решений в различных сферах. Международные и отраслевые стандарты качества данных. Это системы требований и методологий, созданные признанными организациями для обеспечения единого подхода к оценке и улучшению данных. Стандарты в области информационной безопасности. Например, международные и национальные нормы, регламентирующие защиту информации, управление рисками, оценку уязвимостей и методы обеспечения конфиденциальности, целостности и доступности данных
Тема 2.1 Предобработка и подготовка данных. Обработка данных с помощью пакета Pandas	ПК- 5 Способен осуществлять проектирование структур данных баз данных ПК-6 Способен применять методы и средства проектирования компьютерного программного обеспечения	**ИПК-5.1 Знать:** Методы и средства проектирования компьютерного программного обеспечения **ИПК-5.2 Уметь:** Применять методы и средства проектирования компьютерного программного обеспечения, структур данных, баз данных, программных интерфейсов **ИПК-6.1 Знать:** Методы и средства проектирования баз	Самостоятельно изучите тему «Основные статистические методы библиотеки Pandas.» и подготовьте краткий доклад по изученному материалу Что следует отразить: Можно рассмотреть методы расчёта описательных статистик, примеры кода и методы визуализации результатов статистического анализа данных. Также можно изучить ресурсы, которые помогают изучить эти темы. Методы Метод describe() — генерирует сводку описательных статистик для

Тема	Код и формулировка компетенции	Индикаторы достижения компетенции	Типовое задание для самостоятельной работы
		данных **ИПК-6.2 Уметь:** Применять методы и средства проектирования компьютерного программного обеспечения, структур данных, баз данных, программных интерфейсов	всех числовых колонок в DataFrame. Включает количество, среднее, стандартное отклонение, минимум, максимум и значения квартилей. Методы mean() и median() — вычисляют среднее значение числовой колонки и медиану (среднее значение набора данных, отсортированного в порядке возрастания). Методы min() и max() — находят минимальное и максимальное значения в колонке. Метод count() — подсчитывает уникальные значения в категориальной колонке. Функция quantile() — вычисляет квантили (значения, соответствующие p-ому проценту). Примеры кода для реализации статистических методов с помощью Pandas, например: Вычисление среднего значения колонки score — score_mean = df['score'].mean(). Вычисление медианы колонки score — score_median = df['score'].median(). Поиск минимальных и максимальных значений — score_min = df['score'].min() и score_max = df['score'].max(). Визуализация Встроенные функции для создания графиков — метод plot, который является оболочкой вокруг библиотеки matplotlib. Позволяет создавать визуализации напрямую из DataFrame или Series без необходимости извлекать данные и настраивать графики с нуля. Типы графиков — линейный, столбчатый, гистограмма, диаграмма рассеяния, ящик с усами (box). Настройки внешнего вида графиков — Pandas позволяет настраивать внешний вид графиков через параметры метода plot.

Тема	Код и формулировка компетенции	Индикаторы достижения компетенции	Типовое задание для самостоятельной работы
			Примеры графиков и диаграмм, построенных с помощью библиотеки Pandas для визуализации результатов статистического анализа данных: Ресурсы Для изучения статистических методов с помощью Pandas можно использовать: Документацию Pandas — она содержит информацию о методах и функциях библиотеки. Специализированные книги и онлайн-курсы — они могут предоставить больше информации о контексте и применении функций Pandas. Статьи — например, «Описательная статистика Pandas: Руководство для начинающих» на сайте LabEx, где приведены пошаговые инструкции по расчёту статистик.

Критерии оценивания самостоятельной работы:

Оценка	Критерии оценивания
отлично	выставляется если работа носит научно-исследовательский характер, проанализирован и сделан сравнительный анализ нескольких литературных источников, приведены примеры
хорошо	выставляется если проанализирован и сделан сравнительный анализ нескольких литературных источников, приведены примеры
удовлетворительно	выставляется если проведен сравнительный анализ научно-методической литературы, приведены примеры
неудовлетворительно	выставляется если работа прошла проверку на антиплагиат и соответствует требованиям оформления

4.5. ИНДИВИДУАЛЬНЫЕ ЗАДАНИЯ ДЛЯ КУРСОВЫХ РАБОТ

Курсовые работы не предусмотрены

Курсовая работа позволяет выявить степень владения базовыми знаниями, умениями и навыками, необходимыми для обучения, и определить уровень владения новым материалом.

Примерные индивидуальные задания (темы) для курсовых работ:

Критерии оценивания курсовой работы:

Оценка	Критерии оценивания
отлично	Содержание курсовой работы полностью соответствует заданию, содержащемуся в методических указаниях, и плану. Представлены результаты структурированного и логически последовательного обзора литературных и иных источников по теме исследования. Структура курсовой работы логически и методически выдержана. Верно определены исходные данные для расчетов. Все аналитические расчеты выполнены верно, корректно применены методы экономического анализа, не нарушена методика анализа предмета исследования. Все выводы и предложения убедительно аргументированы. При защите курсовой работы обучающийся правильно и уверенно отвечает на вопросы преподавателя, демонстрирует глубокое знание теоретического материала, способен аргументировать собственные утверждения и выводы
хорошо	Содержание курсовой работы полностью соответствует заданию, содержащемуся в методических указаниях, и плану. Представлены результаты структурированного и логически последовательного обзора литературных и иных источников по теме исследования. Структура курсовой работы логически и методически выдержана. Верно определены исходные данные для расчетов. В расчетах допускаются незначительные (не искажающие общего итога оценки) погрешности/ошибки. Большинство выводов и предложений аргументировано, корректно применены методы экономического анализа, не нарушена методика анализа предмета исследования. Оформление курсовой работы и полученные результаты в целом отвечают требованиям, изложенным в методических указаниях. Имеются одна-две несущественные ошибки в использовании терминов, в построенных диаграммах и схемах, в оформлении таблиц. Наличествует незначительное количество грамматических и/или стилистических ошибок. При защите курсовой работы обучающийся правильно и уверенно отвечает на большинство вопросов преподавателя, демонстрирует хорошее знание теоретического материала, но не всегда способен аргументировать собственные утверждения и выводы. При наводящих вопросах преподавателя исправляет ошибки в ответе
удовлетворительно	Содержание курсовой работы полностью соответствует заданию, содержащемуся в методических указаниях, и плану. Результаты обзора литературных и иных источников представлены недостаточно полно, недостаточно логично и последовательно. Верно определены исходные данные для расчетов, но имеются грубые ошибки в расчетах. Аргументация выводов и предложений слабая или отсутствует. Экономические выводы носят констатирующий (описательный) характер. Имеются одно-два существенных отклонений от требований в оформлении курсовой работы. Полученные результаты в целом отвечают требованиям, изложенным в методических указаниях. Имеются одна-две существенных ошибки в использовании терминов, в построенных диаграммах и схемах. Много грамматических и/или стилистических ошибок. При защите курсовой работы обучающийся допускает грубые ошибки при ответах на вопросы преподавателя, демонстрирует слабое знание теоретического материала, в большинстве случаев не способен уверенно аргументировать собственные утверждения и выводы
неудовлетворительно	Содержание курсовой работы не соответствует заданию, содержащемуся в методических указаниях, и плану. Неверно определены исходные данные для расчетов, неверно и не корректно применены методы экономического анализа. Экономические выводы содержат неверную экономическую оценку. Имеются более двух существенных отклонений от требований в оформлении курсовой работы. Большое количество существенных ошибок по сути работы, много грамматических и стилистических ошибок и др. Полученные результаты не отвечают требованиям, изложенным в методических указаниях. При защите курсовой работы обучающийся демонстрирует слабое понимание программного материала, студент не может защитить свои решения, допускает грубые фактические ошибки при ответах на поставленные

	вопросы или вовсе не отвечает на них. Курсовая работа не представлена преподавателю. Обучающийся не явился на защиту курсовой работы

5. МАТЕРИАЛЫ ДЛЯ ОЦЕНКИ РЕЗУЛЬТАТОВ ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ

Промежуточная аттестация проводится в форме зачета/зачета с оценкой/экзамена.

5.1. Вопросы к зачету:

1. Опишите процесс анализа данных и основные этапы, которые он включает.
2. Почему анализ данных рассматривается как инструмент поддержки принятия решений, а не как самостоятельная цель?
3. Раскройте смысл информационной пирамиды и охарактеризуйте каждый её уровень.
4. Почему переход от данных к знаниям невозможен без контекста и интерпретации?
5. Какова роль стандартов в процессе анализа данных и их значение для практики.
6. Опишите основные этапы стандарта CRISP-DM и их назначение.
7. Сравните структуры данных Series и DataFrame.
8. Почему DataFrame считается базовой структурой данных при анализе данных в Pandas?
9. Какие задачи решаются на этапе обработки и подготовки данных в Pandas?
10. Почему этап предобработки данных считается одним из самых важных в анализе данных?
11. Как предварительный анализ структуры данных помогает выбрать дальнейшую стратегию обработки данных?
12. Опишите роль методов info(), head() и describe() при первичном анализе DataFrame.
13. Зачем используется визуализация данных на этапе анализа и подготовки данных?
14. Какие преимущества даёт использование встроенных средств визуализации Pandas?
15. Встроенные средства Pandas упрощают и ускоряют визуализацию данных, так как не требуют отдельной настройки и напрямую работают с объектами DataFrame и Series.

Критерии оценивания зачета с оценкой/экзамена:

Оценка	Критерии оценивания
зачтено	Обучающийся должен: уметь строить ответ в соответствии со структурой излагаемого вопроса; продемонстрировать прочное, достаточно полное усвоение знаний программного материала; продемонстрировать знание основных теоретических понятий; правильно формулировать определения; последовательно, грамотно и логически стройно изложить теоретический материал; продемонстрировать умения самостоятельной работы с литературой; уметь сделать достаточно обоснованные выводы по излагаемому материалу.

не зачтено	Обучающийся демонстрирует: незнание значительной части программного материала; не владение понятийным аппаратом дисциплины; существенные ошибки при изложении учебного материала; неумение строить ответ в соответствии со структурой излагаемого вопроса; неумение делать выводы по излагаемому материалу.

Критерии оценивания зачета с оценкой/экзамена:

Оценка	Критерии оценивания
отлично	Выставляется, если обучающимся правильно и полностью раскрыто содержание материала в пределах программы, чётко и правильно даны определения и раскрыто содержание понятий, точно использованы научные и технические термины, в ответе использованы ранее приобретённые теоретические знания, сделаны необходимые выводы и обобщения
хорошо	Выставляется, если обучающимся раскрыто основное содержание материала в пределах программы, даны определения и раскрыто содержание понятий, в ответе использованы ранее приобретённые теоретические знания, сделаны необходимые выводы и обобщения, но присутствуют незначительные нарушения в последовательности изложения, имеются одна-две неточности в содержании ответа
удовлетворительно	Выставляется, если обучающимся содержание учебного материала изложено фрагментарно, не всегда последовательно, не даны определения, не раскрыто содержание понятий, или они изложены с ошибками, допускаются ошибки и неточности в использовании научной терминологии, отсутствуют выводы и обобщения из предыдущего материала, или возможны ошибки в их изложении
неудовлетворительно	Выставляется, если обучающимся основное содержание учебного материала не раскрыто, не даются ответы на основные вопросы, допускаются грубые ошибки в определении понятий, в использовании терминологии, отсутствуют выводы и обобщения

6. МАТЕРИАЛЫ ДЛЯ ДИАГНОСТИЧЕСКОЙ РАБОТЫ

Задания для диагностической работы должны обеспечивать оценку полностью или частично сформированных компетенций. Каждое задание должно быть привязано к тому или иному индикатору сформированности компетенций.

При формировании заданий для диагностической работы необходимо использовать тестовые задания следующих типов:

Тип задания 1. Задания закрытого типа на установление соответствия.

Тип задания 2. Задания закрытого типа на установление последовательности.

Тип задания 3. Задания комбинированного типа, предполагающие выбор одного правильного ответа из предложенных

Тип задания 4. Задания комбинированного типа, предполагающие выбор нескольких ответов из предложенных

Тип задания 5. Задания открытого типа с развернутым ответом.

Все типы заданий должны быть представлены не менее одного раза.

№ п/п	Тема занятия	Код компетенции	Индикатор	Тип задания	Задание		Уровень освоения компетенции
					Вариант 1	Вариант 2	
1	Тема 1.1 Процесс анализа данных.	ПК- 5 Способен осуществлять проектирование структур данных баз данных ПК-6 Способен применять методы и средства проектирования компьютерного программного обеспечения	**ИПК-5.1 Знать:** Методы и средства проектирования компьютерного программного обеспечения **ИПК-5.2 Уметь:** Применять методы и средства проектирования компьютерного программного обеспечения, структур данных, баз данных, программных интерфейсов **ИПК-6.1 Знать:** Методы и средства проектирования баз данных **ИПК-6.2 Уметь:** Применять методы и средства проектирования компьютерного	1,2,3,4,5	**Задание 1** Какие утверждения корректно отражают сущность процесса анализа данных? Варианты ответа (выберите все верные): 1. Анализ данных направлен на извлечение знаний и закономерностей из данных 2. Анализ данных заключается только в хранении данных в базах данных 3. Анализ данных включает этапы сбора, очистки, обработки и интерпретации данных 4. Анализ данных не предполагает формулировки цели исследования 5. Анализ данных используется для поддержки принятия управленческих и прикладных решений Ответ: 1, 3, 5 Пояснение: 1 — верно, основной целью анализа данных является получение знаний и выводов. 3 — верно, анализ данных	**Задание 1** Какие утверждения корректно характеризуют уровни информации и стандарты анализа данных? Варианты ответа (выберите все верные): 1. Данные представляют собой необработанные факты и наблюдения 2. Информация формируется в результате интерпретации данных в контексте 3. Знания находятся на нижнем уровне информационной пирамиды 4. Стандарты анализа данных обеспечивают воспроизводимость и единый подход к анализу 5. CRISP-DM является примером стандарта процесса анализа данных Ответ: 1, 2, 4, 5 Пояснение: 1 — верно, данные не несут смысла без обработки. 2 — верно, информация — это осмысленные данные. 4 — верно, стандарты регламентируют процесс анализа. 5 — верно, CRISP-DM —	**Базовый (1-3 минуты)**

			программного обеспечения, структур данных, баз данных, программных интерфейсов		представляет собой многоэтапный процесс. 5 — верно, результаты анализа используются для принятия решений. **Задание 2** Расположите этапы анализа данных в логически корректной последовательности. Этапы: 1. Моделирование 2. Понимание задачи (целей анализа) 3. Подготовка данных 4. Понимание и первичный анализ данных 5. Оценка результатов 6. Внедрение результатов Ответ: 243156 Пояснение: Сначала формулируется цель анализа, затем изучаются данные, после чего они подготавливаются. Далее строятся модели, оценивается их качество и только после этого результаты внедряются.	общепринятый стандарт жизненного цикла анализа данных. **Задание 2** Расположите уровни информационной пирамиды от базового к наиболее высокому. Уровни пирамиды: 1. Знания 2. Данные 3. Мудрость 4. Информация Ответ: 2413 Пояснение: Данные являются основой, из них формируется информация, затем — знания, а высший уровень — мудрость как способность принимать обоснованные решения. **Задание 3** Установите соответствие между стандартом и его характеристикой. Список 1: 1. CRISP-DM 2. KDD 3. SEMMA Список 2: A. Стандарт, ориентированный на этапы выборки, моделирования и оценки B. Процесс извлечения знаний из	

					Задание 3 Установите соответствие между уровнем информационной пирамиды и его описанием. Список 1: 1. Данные 2. Информация 3. Знания 4. Мудрость Список 2: A. Осмысленные выводы и рекомендации для принятия решений B. Необработанные факты и значения C. Выявленные закономерности и обобщения D. Данные, интерпретированные в определённом контексте Ответ: 1B, 2D, 3C, 4A **Задание 4** Какова основная цель анализа данных? Варианты ответа: 1. Хранение больших объёмов информации 2. Извлечение знаний и поддержка принятия решений 3. Создание информационных систем	баз данных C. Универсальная модель жизненного цикла анализа данных Ответ: 1C, 2B, 3A **Задание 4** Какой стандарт наиболее широко используется для описания жизненного цикла анализа данных? Варианты ответа: 1. UML 2. SQL 3. CRISP-DM 4. HTTP Ответ: 3 Пояснение: CRISP-DM — общепринятый стандарт, описывающий этапы анализа данных от постановки задачи до внедрения. **Задание 5** Как называется стандарт процесса анализа данных, включающий этапы понимания задачи, понимания данных, подготовки данных, моделирования, оценки и внедрения? Ответ: CRISP-DM Пояснение: CRISP-DM описывает полный	

					4. Замена экспертных оценок без проверки Ответ: 2 Пояснение: Анализ данных направлен на получение знаний и обоснование решений, а не на хранение или разработку систем. **Задание 5** Как называется уровень информационной пирамиды, на котором данные приобретают смысл в контексте? Ответ: Информация Пояснение: Информация формируется в результате интерпретации данных и добавления контекста.	жизненный цикл анализа данных и широко применяется в академической и прикладной практике.	
2	Тема 2.1 Предобработка и подготовка данных. Обработка данных с помощью пакета Pandas	ПК- 5 Способен осуществлять проектирование структур данных баз данных	**ИПК-5.1 Знать:** Методы и средства проектирования компьютерного программного обеспечения **ИПК-5.2 Уметь:** Применять методы и средства	Все	**Задание 6** Какие утверждения корректно описывают структуры данных Series и DataFrame библиотеки Pandas? Варианты ответа (выберите все верные): 1. Series представляет собой одномерную структуру данных с индексами	**Задание 6** Какие методы Pandas используются для предварительного анализа структуры данных и визуализации? Варианты ответа (выберите все верные): 1. head() 2. info() 3. describe() 4. plot()	**Повышенный (3-5 минут)**

		ПК-6 Способен применять методы и средства проектирования компьютерного программного обеспечения	проектирования компьютерного программного обеспечения, структур данных, баз данных, программных интерфейсов **ИПК-6.1 Знать:** Методы и средства проектирования баз данных **ИПК-6.2 Уметь:** Применять методы и средства проектирования компьютерного программного обеспечения, структур данных, баз данных, программных интерфейсов		2. DataFrame является двумерной табличной структурой данных 3. Series может содержать данные разных типов в одном объекте 4. DataFrame состоит из набора объектов типа Series 5. Series не поддерживает работу с индексами Ответ: 1, 2, 4 Пояснение: 1 — верно, Series — одномерный массив с индексами. 2 — верно, DataFrame представляет собой таблицу (строки и столбцы). 4 — верно, каждый столбец DataFrame является объектом Series. **Задание 7** Расположите этапы работы с данными в Pandas в логически корректной последовательности. Этапы: 1. Визуальный анализ данных 2. Загрузка данных в DataFrame 3. Анализ структуры данных (info(), head()) 4. Очистка и преобразование данных Ответ: 2341	5. merge() Ответ: 1, 2, 3, 4 Пояснение: 1 — вывод первых строк набора данных. 2 — отображает информацию о типах данных и пропусках. 3 — рассчитывает описательные статистики. 4 — используется для визуализации данных. **Задание 7** Упорядочите действия при первичном анализе нового набора данных в Pandas. Действия: 1. Построение графиков 2. Просмотр первых строк данных 3. Получение информации о типах данных и пропусках 4. Расчёт описательных статистик Ответ: 2341 Пояснение: Сначала изучается содержимое таблицы, затем её структура, далее числовые характеристики, и только после этого — визуализация. **Задание 8** Установите соответствие между структурой данных Pandas и её	

					Пояснение: Сначала данные загружаются, затем изучается их структура, после чего выполняется очистка, и только затем проводится визуальный анализ. **Задание 8** Установите соответствие между методом Pandas и его назначением. Список 1: 1. head() 2. info() 3. describe() 4. plot() Список 2: A. Визуализация данных B. Вывод общей информации о структуре DataFrame C. Отображение первых строк таблицы D. Расчёт описательных статистик Ответ: 1C, 2B, 3D, 4A **Задание 9** Какой метод Pandas позволяет получить информацию о типах данных столбцов и количестве пропусков?	характеристикой. Список 1: 1. Series 2. DataFrame Список 2: A. Двумерная табличная структура данных B. Одномерный массив с индексами Ответ: 1B, 2A **Задание 9** Какой метод Pandas используется для построения графиков на основе данных DataFrame? Варианты ответа: 1. visualize() 2. draw() 3. plot() 4. show() Ответ: 3 Пояснение: Метод plot() предоставляет встроенные средства визуализации на основе Matplotlib. **Задание 10** Какой метод Pandas используется для вычисления основных описательных статистик числовых столбцов?	

					Варианты ответа: 1. head() 2. describe() 3. info() 4. plot() Ответ: 3 Пояснение: Метод info() выводит типы данных, количество ненулевых значений и объём памяти. **Задание 10** Как называется двумерная табличная структура данных библиотеки Pandas? Ответ: DataFrame Пояснение: DataFrame предназначен для хранения и обработки табличных данных с индексами и столбцами.	Ответ: describe() Пояснение: Метод describe() рассчитывает среднее, медиану, стандартное отклонение и другие статистические показатели.	

Критерии оценивания диагностической работы:

Оценка	Критерии оценивания
отлично	от 90 до 100 % правильно выполненных заданий
хорошо	от 70 до 89 % правильно выполненных заданий
удовлетворительно	от 50 до 69 % правильно выполненных заданий
неудовлетворительно	менее 50 % правильно выполненных заданий