

МИНОБРНАУКИ РОССИИ
ВЛАДИВОСТОКСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
КАФЕДРА МАТЕМАТИКИ И МОДЕЛИРОВАНИЯ

Рабочая программа дисциплины (модуля)
АНАЛИЗ ДАННЫХ НА PYTHON

Направление и направленность (профиль)
38.03.05 Бизнес-информатика. Бизнес-аналитика

Год набора на ОПОП
2023

Форма обучения
очная

Владивосток 2025

Рабочая программа дисциплины (модуля) «Анализ данных на Python» составлена в соответствии с требованиями ФГОС ВО по направлению подготовки 38.03.05 Бизнес-информатика (утв. приказом Минобрнауки России от 29.07.2020г. №838) и Порядком организации и осуществления образовательной деятельности по образовательным программам высшего образования – программам бакалавриата, программам специалитета, программам магистратуры (утв. приказом Минобрнауки России от 06.04.2021 г. N245).

Составитель(и):

Красько А.А., кандидат экономических наук, доцент, Кафедра математики и моделирования, Andrey.Krasko@vvsu.ru

Утверждена на заседании кафедры математики и моделирования от 15.05.2025 ,
протокол № 9

СОГЛАСОВАНО:

Заведующий кафедрой (разработчика)

Галимзянова К.Н.

ДОКУМЕНТ ПОДПИСАН ЭЛЕКТРОННОЙ ПОДПИСЬЮ	
Сертификат	1599657997
Номер транзакции	0000000000DE2CA4
Владелец	Галимзянова К.Н.

1 Цель, планируемые результаты обучения по дисциплине (модулю)

Формирование у студентов знаний теоретических основ, практических навыков и умений использования языка программирования Python, владение набором компетенций в области программирования и анализа данных.

Планируемыми результатами обучения по дисциплине (модулю), являются знания, умения, навыки. Перечень планируемых результатов обучения по дисциплине (модулю), соотнесенных с планируемыми результатами освоения образовательной программы, представлен в таблице 1.

Таблица 1 – Компетенции, формируемые в результате изучения дисциплины (модуля)

Название ОПОП ВО, сокращенное	Код и формулировка компетенции	Код и формулировка индикатора достижения компетенции	Результаты обучения по дисциплине		
			Код результата	Формулировка результата	
38.03.05 «Бизнес-информатика» (Б-БИ)	ОПК-3 : Способен управлять процессами создания и использования продуктов и услуг в сфере информационно-коммуникационных технологий, в том числе разрабатывать алгоритмы и программы для их практической реализации	ОПК-3.2к : Обладает умением разрабатывать алгоритмы и программные продукты, регистрировать права на интеллектуальную собственность	РД1	Знание	основ разработки алгоритмов и программных средств, а так же механизмов регистрации прав на интеллектуальную собственность
			РД3	Умение	разрабатывать алгоритмы и программные средства для решения прикладных задач
			РД4	Знание	современных методов, моделей, алгоритмов, технологий и инструментальных средств для работы при решении прикладных задач с использованием больших данных
	ОПК-4 : Способен понимать принципы работы информационных технологий; использовать информацию, методы и программные средства ее сбора, обработки и анализа для информационно-аналитической поддержки принятия управленческих решений	ОПК-4.2к : Использует методы, модели, алгоритмы, технологии и инструментальные средства для работы с большими данными при решении прикладных задач	РД2	Навык	выбора инструментов для проведения анализа организации
			РД5	Умение	выбирать методы, модели, алгоритмы, технологии и инструментальные средства для решения прикладных задач с использованием больших данных
			РД6	Навык	применения современных методов, моделей, алгоритмов, технологий и инструментальных средств для решения прикладных задач с использованием больших данных

В процессе освоения дисциплины решаются задачи воспитания гармонично развитой, патриотичной и социально ответственной личности на основе традиционных

российских духовно-нравственных и культурно-исторических ценностей, представленные в таблице 1.2.

Таблица 1.2 – Целевые ориентиры воспитания

Воспитательные задачи	Формирование ценностей	Целевые ориентиры
1 Формирование гражданской позиции и патриотизма		
Развитие патриотизма и гражданской ответственности	Гражданственность	Гибкость мышления
2 Формирование духовно-нравственных ценностей		
Воспитание чувства долга и ответственности перед семьей и обществом	Гражданственность	Способность находить, анализировать и структурировать информацию
3 Формирование научного мировоззрения и культуры мышления		
Развитие познавательного интереса и стремления к знаниям	Созидательный труд	Любознательность
4 Формирование коммуникативных навыков и культуры общения		
Воспитание культуры диалога и уважения к мнению других людей	Взаимопомощь и взаимоуважение	Внимательность к деталям

2 Место дисциплины (модуля) в структуре ОПОП

Дисциплина относится к обязательной части блока 1 дисциплин учебного плана направлений "Бизнес-информатика. Бизнес-аналитика" и "Прикладная математика. Интеллектуальный анализ данных"

3. Объем дисциплины (модуля)

Объем дисциплины (модуля) в зачетных единицах с указанием количества академических часов, выделенных на контактную работу с обучающимися (по видам учебных занятий) и на самостоятельную работу, приведен в таблице 2.

Таблица 2 – Общая трудоемкость дисциплины

Название ОПОП ВО	Форма обуче- ния	Часть УП	Семестр (ОФО) или курс (ЗФО, ОЗФО)	Трудо- емкость	Объем контактной работы (час)						СРС	Форма аттес- тации
				(З.Е.)	Всего	Аудиторная			Внеауди- торная			
						лек.	прак.	лаб.	ПА	КСР		
38.03.05 Бизнес- информатика	ОФО	Б1.Б	4	3	55	18	36	0	1	0	53	Э

4 Структура и содержание дисциплины (модуля)

4.1 Структура дисциплины (модуля) для ОФО

Тематический план, отражающий содержание дисциплины (перечень разделов и тем), структурированное по видам учебных занятий с указанием их объемов в соответствии с учебным планом, приведен в таблице 3.1

Таблица 3.1 – Разделы дисциплины (модуля), виды учебной деятельности и формы текущего контроля для ОФО

№	Название темы	Код результата обучения	Кол-во часов, отведенное на				Форма текущего контроля
			Лек	Практ	Лаб	СРС	
1	Введение в программирование. Ввод, обработка, вывод данных в Python	РД1, РД3	2	4	0	6	групповая дискуссия по основным вопросам и проблемам изучаемой темы
2	Работа с табличными данными	РД1, РД3	4	6	0	10	групповая дискуссия по основным вопросам и проблемам изучаемой темы
3	Визуализация данных, библиотеки: Matplotlib, Seaborn, plotly	РД1, РД2	2	6	0	7	групповая дискуссия по основным вопросам и проблемам изучаемой темы
4	Элементы статистики, библиотека SciPy	РД5, РД6	4	6	0	10	индивидуальное домашнее задание
5	Машинное обучение Python	РД4, РД5, РД6	4	8	0	10	индивидуальное домашнее задание
6	Кросс-валидация. Логистическая регрессия. Метод k-ближайших соседей	РД4, РД6	2	6	0	10	контрольная работа
Итого по таблице			18	36	0	53	

4.2 Содержание разделов и тем дисциплины (модуля) для ОФО

Тема 1 Введение в программирование. Ввод, обработка, вывод данных в Python.

Содержание темы: Установка языка Python, настройка. Программирование на Python: Объекты и Структуры Данных в Python; Операторы Сравнения в Python; Операторы Python; Методы и Функции.

Формы и методы проведения занятий по теме, применяемые образовательные технологии: лекция.

Виды самостоятельной подготовки студентов по теме: подготовка к дискуссии, изучение лекционного материала.

Тема 2 Работа с табличными данными.

Содержание темы: Преобразование табличных данных, селекция таблиц, библиотека Pandas, объекты Pandas.Series и Pandas.DataFrame. Анализ и исследования данных одной таблицы (работа с единым датафреймом). Работа с несколькими таблицами, операции merge, join, различные типы join'ов. Основы sql-синтаксиса, библиотека pandas.

Формы и методы проведения занятий по теме, применяемые образовательные технологии: лекция.

Виды самостоятельной подготовки студентов по теме: подготовка к дискуссии, изучение лекционного материала.

Тема 3 Визуализация данных, библиотеки: Matplotlib, Seaborn, plotly.

Содержание темы: Рассмотрение функционала и областей применения библиотек визуализации данных: Matplotlib, Seaborn, plotly.

Формы и методы проведения занятий по теме, применяемые образовательные технологии: лекция.

Виды самостоятельной подготовки студентов по теме: подготовка к дискуссии, изучение лекционного материала.

Тема 4 Элементы статистики, библиотека SciPy.

Содержание темы: Основные вероятностные распределения. Нормальное распределение. Центральная предельная теорема и визуализация ее утверждения методами Python. Доверительные интервалы. Сравнение двух средних, t-критерий Стьюдента. Проверка распределения на нормальность.

Формы и методы проведения занятий по теме, применяемые образовательные технологии: лекция, практическое занятие, метод кооперативного обучения.

Виды самостоятельной подготовки студентов по теме: подготовка к практическому занятию, выполнение индивидуального домашнего задания, изучение лекционного материала.

Тема 5 Машинное обучение Python.

Содержание темы: Машинное обучение: Линейная Регрессия; Полиномиальная регрессия; Регуляризация. Библиотека Scikit-Learn Конструирование признаков (Feature Engineering) и подготовка данных.

Формы и методы проведения занятий по теме, применяемые образовательные технологии: лекция, практическое занятие, метод кооперативного обучения.

Виды самостоятельной подготовки студентов по теме: подготовка к практическому занятию, выполнение индивидуального домашнего задания, изучение лекционного материала.

Тема 6 Кросс-валидация. Логистическая регрессия. Метод k-ближайших соседей.

Содержание темы: Корсс-валидация: цели, задачи, алгоритмы. Логистическая регрессия: виды, метрики качества построенных моделей. Метод k-ближайших соседей.

Формы и методы проведения занятий по теме, применяемые образовательные технологии: лекция, практическое занятие, метод кооперативного обучения.

Виды самостоятельной подготовки студентов по теме: подготовка к контрольной работе, подготовка к практическому занятию, изучение лекционного материала.

5 Методические указания для обучающихся по изучению и реализации дисциплины (модуля)

5.1 Методические рекомендации обучающимся по изучению дисциплины и по обеспечению самостоятельной работы

Для лучшего освоения материала и систематизации знаний по дисциплине необходимо постоянно разбирать материалы лекций. Во время самостоятельной проработки лекционного материала особое внимание следует уделять возникшим вопросам, непонятным терминам, спорным точкам зрения. В случае необходимости обращаться к преподавателю за консультацией. Студент должен четко уяснить, что именно с лекции начинается его подготовка к практическому занятию. Вместе с тем, лекция лишь организует мыслительную деятельность, но не обеспечивает глубину усвоения программного материала.

При подготовке к практическому занятию особое внимание необходимо обратить на содержание основных положений и выводов, объяснение явлений и фактов, уяснение практического приложения рассматриваемых теоретических вопросов. В процессе этой работы студент должен стремиться понять и запомнить основные положения рассматриваемого материала, примеры, поясняющие его, а также разобраться в иллюстративном материале. В процессе подготовки к практическому занятию рекомендуется взаимное обсуждение материала, во время которого закрепляются знания, а также приобретается практика в изложении и разъяснении полученных знаний.

При выполнении индивидуальных домашних заданий необходимо использовать теоретический материал, делать ссылки на соответствующие теоремы, свойства, формулы и др. Решение выполняется подробно и содержит необходимые пояснительные ссылки.

Самостоятельная работа также включает работу на практических занятиях, во время применения «Метода кооперативного обучения» студенты работают в малых группах (3 – 4 чел.) над заданиями, в процессе выполнения которых они могут совещаться друг с другом, а также обращаться за помощью к преподавателю.

5.2 Особенности организации обучения для лиц с ограниченными возможностями здоровья и инвалидов

При необходимости обучающимся из числа лиц с ограниченными возможностями здоровья и инвалидов (по заявлению обучающегося) предоставляется учебная информация в доступных формах с учетом их индивидуальных психофизических особенностей:

- для лиц с нарушениями зрения: в печатной форме увеличенным шрифтом; в форме электронного документа; индивидуальные консультации с привлечением тифлосурдопереводчика; индивидуальные задания, консультации и др.

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

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

6 Фонд оценочных средств для проведения текущего контроля и промежуточной аттестации обучающихся по дисциплине (модулю)

В соответствии с требованиями ФГОС ВО для аттестации обучающихся на соответствие их персональных достижений планируемым результатам обучения по дисциплине (модулю) созданы фонды оценочных средств. Типовые контрольные задания, методические материалы, определяющие процедуры оценивания знаний, умений и навыков, а также критерии и показатели, необходимые для оценки знаний, умений, навыков и характеризующие этапы формирования компетенций в процессе освоения образовательной программы, представлены в Приложении 1.

7 Учебно-методическое и информационное обеспечение дисциплины (модуля)

7.1 Основная литература

1. Коротеев, М. В., Основы машинного обучения на Python : учебник / М. В. Коротеев. — Москва : КноРус, 2024. — 431 с. — ISBN 978-5-406-12673-8. — URL: <https://book.ru/book/952751> (дата обращения: 18.06.2025). — Текст : электронный.

2. Чернышев, С. А., Алгоритмы и структуры данных на Python : учебное пособие / С. А. Чернышев. — Москва : КноРус, 2024. — 326 с. — ISBN 978-5-406-11683-8. — URL: <https://book.ru/book/949701> (дата обращения: 18.06.2025). — Текст : электронный.

7.2 *Дополнительная литература*

1. Платонов, А. В. Машинное обучение : учебное пособие для вузов / А. В. Платонов. — Москва : Издательство Юрайт, 2024. — 85 с. — (Высшее образование). — ISBN 978-5-534-15561-7. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/544780> (дата обращения: 12.03.2025).

7.3 *Ресурсы информационно-телекоммуникационной сети "Интернет", включая профессиональные базы данных и информационно-справочные системы (при необходимости):*

1. Научная электронная библиотека «eLIBRARY. RU» [Электронный ресурс] // Режим доступа: <http://elibrary.ru/defaultx.asp>
2. Образовательная платформа "ЮРАЙТ"
3. Электронно-библиотечная система "BOOK.ru"
4. Open Academic Journals Index (ОАИ). Профессиональная база данных - Режим доступа: <http://oaji.net/>
5. Президентская библиотека им. Б.Н.Ельцина (база данных различных профессиональных областей) - Режим доступа: <https://www.prilib.ru/>
6. Информационно-справочная система "Консультант Плюс" - Режим доступа: <http://www.consultant.ru/>

8 Материально-техническое обеспечение дисциплины (модуля) и перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), включая перечень программного обеспечения

Основное оборудование:

- Компьютеры
- Проектор
- Система аудиовизуального представления информации

Программное обеспечение:

- □ Microsoft Office 2010 Standart
- □ Python

МИНОБРНАУКИ РОССИИ
ВЛАДИВОСТОКСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
КАФЕДРА МАТЕМАТИКИ И МОДЕЛИРОВАНИЯ

Фонд оценочных средств
для проведения текущего контроля
и промежуточной аттестации по дисциплине (модулю)

АНАЛИЗ ДАННЫХ НА PYTHON

Направление и направленность (профиль)
38.03.05 Бизнес-информатика. Бизнес-аналитика

Год набора на ОПОП
2023

Форма обучения
очная

Владивосток 2025

1 Перечень формируемых компетенций

Название ОПОП ВО, сокращенное	Код и формулировка компетенции	Код и формулировка индикатора достижения компетенции
38.03.05 «Бизнес-информатика» (Б-БИ)	ОПК-3 : Способен управлять процессами создания и использования продуктов и услуг в сфере информационно-коммуникационных технологий, в том числе разрабатывать алгоритмы и программы для их практической реализации	ОПК-3.2к : Обладает умением разрабатывать алгоритмы и программные продукты, регистрировать права на интеллектуальную собственность
	ОПК-4 : Способен понимать принципы работы информационных технологий; использовать информацию, методы и программные средства ее сбора, обработки и анализа для информационно-аналитической поддержки принятия управленческих решений	ОПК-4.2к : Использует методы, модели, алгоритмы, технологии и инструментальные средства для работы с большими данными при решении прикладных задач

Компетенция считается сформированной на данном этапе в случае, если полученные результаты обучения по дисциплине оценены положительно (диапазон критериев оценивания результатов обучения «зачтено», «удовлетворительно», «хорошо», «отлично»). В случае отсутствия положительной оценки компетенция на данном этапе считается несформированной.

2 Показатели оценивания планируемых результатов обучения

Компетенция ОПК-3 «Способен управлять процессами создания и использования продуктов и услуг в сфере информационно-коммуникационных технологий, в том числе разрабатывать алгоритмы и программы для их практической реализации»

Таблица 2.1 – Критерии оценки индикаторов достижения компетенции

Код и формулировка индикатора достижения компетенции	Результаты обучения по дисциплине			Критерии оценивания результатов обучения
	Код	Тип	Результат	
ОПК-3.2к : Обладает умением разрабатывать алгоритмы и программные продукты, регистрировать права на интеллектуальную собственность	РД 1	Знание	основ разработки алгоритмов и программных средств, а также механизмов регистрации прав на интеллектуальную собственность	правильность ответов на поставленные вопросы, правильность формулировки и анализа проблем
	РД 3	Умение	разрабатывать алгоритмы и программные средства для решения прикладных задач	корректность выбора инструментов решения задач, выполнение всех необходимых расчетов
	РД 4	Знание	современных методов, моделей, алгоритмов, технологий и инструментальных средств для работы при решении прикладных задач с использованием больших данных	правильность ответов на поставленные вопросы, правильность формулировки и анализа проблем

Компетенция ОПК-4 «Способен понимать принципы работы информационных технологий; использовать информацию, методы и программные средства ее сбора, обработки и анализа для информационно-аналитической поддержки принятия управленческих решений»

Таблица 2.2 – Критерии оценки индикаторов достижения компетенции

Код и формулировка индикатора достижения компетенции	Результаты обучения по дисциплине			Критерии оценивания результатов обучения
	Код	Тип	Результат	
ОПК-4.2к : Использует методы, модели, алгоритмы, технологии и инструментальные средства для работы с большими данными при решении прикладных задач	РД 2	Навык	выбора инструментов для проведения анализа организации	корректность выбора инструментов решения задач, выполнение всех необходимых расчетов
	РД 5	Умение	выбирать методы, модели, алгоритмы, технологии и инструментальные средства для решения прикладных задач с использованием больших данных	корректность выбора инструментов решения задач, выполнение всех необходимых расчетов
	РД 6	Навык	применения современных методов, моделей, алгоритмов, технологий и инструментальных средств для решения прикладных задач с использованием больших данных	правильное использование алгоритма выполнения действий, самостоятельность решения поставленных задач

Таблица заполняется в соответствии с разделом 1 Рабочей программы дисциплины (модуля).

3 Перечень оценочных средств

Таблица 3 – Перечень оценочных средств по дисциплине (модулю)

Контролируемые планируемые результаты обучения		Контролируемые темы дисциплины	Наименование оценочного средства и представление его в ФОС	
			Текущий контроль	Промежуточная аттестация
Очная форма обучения				
РД1	Знание : основ разработки алгоритмов и программных средств, а так же механизмов регистрации и прав на интеллектуальную собственность	1.1. Введение в программирование. Ввод, обработка, вывод данных в Python	Тест	Тест
		1.2. Работа с табличными данными	Тест	Тест
		1.3. Визуализация данных, библиотеки: Matplotlib, Seaborn, plotly	Тест	Тест
РД2	Навык : выбора инструментов для проведения анализа организации	1.3. Визуализация данных, библиотеки: Matplotlib, Seaborn, plotly	Тест	Тест

РД3	Умение : разрабатывать алгоритмы и программные средства для решения прикладных задач	1.1. Введение в программирование. Ввод, обработка, вывод данных в Python	Разноуровневые задачи и задания	Зачет в письменной форме
			Разноуровневые задачи и задания	Тест
			Тест	Зачет в письменной форме
			Тест	Тест
		1.2. Работа с табличным и данными	Разноуровневые задачи и задания	Зачет в письменной форме
			Разноуровневые задачи и задания	Тест
			Тест	Зачет в письменной форме
			Тест	Тест
РД4	Знание : современных методов, моделей, алгоритмов, технологий и инструментальных средств для работы при решении прикладных задач с использованием больших данных	1.5. Машинное обучение Python	Разноуровневые задачи и задания	Зачет в письменной форме
			Разноуровневые задачи и задания	Тест
			Тест	Зачет в письменной форме
			Тест	Тест
		1.6. Кросс-валидация. Логистическая регрессия. Метод k-ближайших соседей	Разноуровневые задачи и задания	Зачет в письменной форме
			Разноуровневые задачи и задания	Тест
			Тест	Зачет в письменной форме
			Тест	Тест
РД5	Умение : выбирать методы, модели, алгоритмы, технологии и инструментальные средства для решения прикладных задач с использованием больших данных	1.4. Элементы статистики, библиотека SciPy	Разноуровневые задачи и задания	Зачет в письменной форме
			Разноуровневые задачи и задания	Тест
			Тест	Зачет в письменной форме
			Тест	Тест
		1.5. Машинное обучение Python	Разноуровневые задачи и задания	Зачет в письменной форме
			Разноуровневые задачи и задания	Тест
			Тест	Зачет в письменной форме
			Тест	Тест

РД6	Навык : применения современных методов, моделей, алгоритмов, технологий и инструментальных средств для решения прикладных задач с использованием больших данных	1.4. Элементы статистики, библиотека SciPy	Разноуровневые задачи и задания	Зачет в письменной форме
			Разноуровневые задачи и задания	Тест
			Тест	Зачет в письменной форме
			Тест	Тест
		1.5. Машинное обучение Python	Разноуровневые задачи и задания	Зачет в письменной форме
			Разноуровневые задачи и задания	Тест
			Тест	Зачет в письменной форме
			Тест	Тест
		1.6. Кросс-валидация. Логистическая регрессия. Метод k-ближайших соседей	Разноуровневые задачи и задания	Зачет в письменной форме

4 Описание процедуры оценивания

Качество сформированности компетенций на данном этапе оценивается по результатам текущих и промежуточных аттестаций при помощи количественной оценки, выраженной в баллах. Максимальная сумма баллов по дисциплине (модулю) равна 100 баллам.

Виды оценивания	Разноуровневые задачи и задания	Тест	Итого
Лекция		40	40
Практика	50		50
Промежуточная аттестация		10	10
Итого	50	50	100

Сумма баллов, набранных студентом по всем видам учебной деятельности в рамках дисциплины, переводится в оценку в соответствии с таблицей.

Сумма баллов по дисциплине	Оценка по промежуточной аттестации	Характеристика качества сформированности компетенции
от 91 до 100	«зачтено» / «ОТЛИЧНО»	Студент демонстрирует сформированность дисциплинарных компетенций, обнаруживает всестороннее, систематическое и глубокое знание учебного материала, усвоил основную литературу и знаком с дополнительной литературой, рекомендованной программой, умеет свободно выполнять практические задания, предусмотренные программой, свободно оперирует приобретенными знаниями, умениями, применяет их в ситуациях повышенной сложности.
от 76 до 90	«зачтено» / «хорошо»	Студент демонстрирует сформированность дисциплинарных компетенций: основные знания, умения освоены, но допускаются незначительные ошибки, неточности, затруднения при аналитических операциях, переносе знаний и умений на новые, нестандартные ситуации.
от 61 до 75	«зачтено» /	Студент демонстрирует сформированность дисциплинарных компетенций: в ходе контрольных мероприятий допускаются значительные ошибки, проявляется отсутствие отдельных знаний, умений, навыков по некоторым дисциплинарным

	«удовлетворительно»	м компетенциям, студент испытывает значительные затруднения при оперировании знаниями и умениями при их переносе на новые ситуации.
от 41 до 60	«не зачтено» / «неудовлетворительно»	У студента не сформированы дисциплинарные компетенции, проявляется недостаточность знаний, умений, навыков.
от 0 до 40	«не зачтено» / «неудовлетворительно»	Дисциплинарные компетенции не сформированы. Проявляется полное или практически полное отсутствие знаний, умений, навыков.

Сумма баллов, набранных студентом по всем видам учебной деятельности в рамках дисциплины, переводится в оценку в соответствии с таблицей.

Сумма баллов по дисциплине	Оценка по промежуточной аттестации	Характеристика качества сформированности компетенции
от 91 до 100	«зачтено» / «отлично»	Студент демонстрирует сформированность дисциплинарных компетенций, обладает всестороннее, систематическое и глубокое знание учебного материала, усвоил основную литературу и знаком с дополнительной литературой, рекомендованной программой, умеет свободно выполнять практические задания, предусмотренные программой, свободно оперирует приобретенными знаниями, умениями, применяет их в ситуациях повышенной сложности.
от 76 до 90	«зачтено» / «хорошо»	Студент демонстрирует сформированность дисциплинарных компетенций: основные знания, умения освоены, но допускаются незначительные ошибки, неточности, затруднения при аналитических операциях, переносе знаний и умений на новые, нестандартные ситуации.
от 61 до 75	«зачтено» / «удовлетворительно»	Студент демонстрирует сформированность дисциплинарных компетенций: в ходе контрольных мероприятий допускаются значительные ошибки, проявляется отсутствие отдельных знаний, умений, навыков по некоторым дисциплинарным компетенциям, студент испытывает значительные затруднения при оперировании знаниями и умениями при их переносе на новые ситуации.
от 41 до 60	«не зачтено» / «неудовлетворительно»	У студента не сформированы дисциплинарные компетенции, проявляется недостаточность знаний, умений, навыков.
от 0 до 40	«не зачтено» / «неудовлетворительно»	Дисциплинарные компетенции не сформированы. Проявляется полное или практически полное отсутствие знаний, умений, навыков.

5 Примерные оценочные средства

5.1 Пример разноуровневых задач и заданий

Приводятся графики изменения объемов инвестиций и прибыли за 10-15 лет по 4 предприятиям.

Провести раздельное оценивание уравнений регрессии. Построить модели с фиксированными и случайными эффектами. Оценить качество моделей, сравнить результаты. Применить критерий оценки Хаусмана.

Краткие методические указания

При выполнении разноуровневых задач и заданий особое внимание необходимо обратить на содержание основных положений и выводов, объяснение явлений и фактов, уяснение практического приложения рассматриваемых теоретических вопросов по соответствующей теме. В процессе этой работы студент должен стремиться понять и запомнить основные положения рассматриваемого материала, примеры, поясняющие его, а также разобраться в иллюстративном материале.

Шкала оценки

Шкала оценки

Оценка	Баллы	Описание
5	40–50	выставляется студенту, если он выполнил без существенных ошибок все задания и ответил на все поставленные вопросы, подтверждая знание материала, умение использовать нормативные документы для подтверждения правильности собственной позиции;
4	25–39	выставляется студенту, если выполнил без существенных ошибок больше половины заданий и ответил на большинство поставленных вопросы, четко представлял свою позицию, подтверждая знание материала, умение использовать нормативные документы для подтверждения правильности собственной позиции;
3	16–24	выставляется студенту, если выполнил без существенных ошибок меньше половины заданий, ответил на некоторые поставленные вопросы, подтверждая знание материала, умение использовать нормативные документы для подтверждения правильности собственной позиции;
2	7–15	выставляется студенту, если он допустил ошибки при ответах на все из поставленных в задаче вопросов.

5.2 Контрольный тест

1. Выберите правильный способ создания пустого списка в Python: a) `list = []` b) `list = list()` c) `list = new List()` d) `list = {empty}`
2. Установите соответствие между типами данных Python и их описанием:
 1. `int`
 2. `float`
 3. `str`
 4. `bool`

A) Целочисленный тип B) Строковый тип C) Логический тип D) Тип с плавающей точкой
3. Напишите название метода списка в Python, который добавляет элемент в конец списка: _____
4. Выберите все корректные способы создания словаря в Python: a) `dict = {}` b) `dict = dict()` c) `dict = [:]` d) `dict = {1: "one", 2: "two"}`
5. Верно ли утверждение: "В Python строки являются неизменяемым (immutable) типом данных"?
6. Какой оператор используется для возведения в степень в Python? a) `^` b) `**` c) `pow()` d) `^^`
7. Установите соответствие между функциями и их назначением:
 1. `len()`
 2. `type()`
 3. `range()`
 4. `print()`

A) Вывод данных на экран B) Определение длины объекта C) Создание последовательности чисел D) Определение типа объекта
8. Напишите название метода строки, который преобразует все символы в нижний регистр: _____
9. Выберите все встроенные структуры данных в Python: a) Список b) Кортеж c) Множество d) Массив e) Словарь
10. Верно ли утверждение: "В Python индексация элементов списка начинается с 1"?
11. Какая библиотека используется для создания базовых графиков в Python? a) Seaborn b) Matplotlib c) Plotly d) Bokeh
12. Установите соответствие между типами графиков и их назначением:
 1. `scatter`
 2. `bar`
 3. `pie`
 4. `line`

А) Линейный график В) Круговая диаграмма С) Точечная диаграмма D) Столбчатая диаграмма

13. Напишите название метода библиотеки matplotlib для отображения графика: _____

14. Выберите все параметры, которые можно настроить в графике matplotlib: а) Цвет линии б) Толщина линии с) Стилль маркероов d) Размер фигуры

15. Верно ли утверждение: "Библиотека Seaborn построена поверх matplotlib"?

16. Какая функция используется для создания подграфиков в matplotlib? а) subplot() б) subplots() с) create_subplot() d) make_subplot()

17. Установите соответствие между библиотеками визуализации и их особенностями:

1. Matplotlib
2. Seaborn
3. Plotly
4. Bokeh

А) Интерактивные графики В) Базовая библиотека С) Статистическая визуализация D) Веб-ориентированная визуализация

18. Напишите название параметра в seaborn, отвечающего за стиль графика: _____

19. Выберите все типы графиков, доступные в библиотеке seaborn: а) heatmap б) boxplot с) violinplot d) swarmplot

20. Верно ли утверждение: "Plotly позволяет создавать только статические графики"?

21. Какая библиотека используется для машинного обучения в Python? а) NumPy б) Pandas с) Scikit-learn d) Matplotlib

22. Установите соответствие между методами регрессии и их применением:

1. LinearRegression
2. Ridge
3. Lasso
4. ElasticNet

А) Линейная регрессия с L2 регуляризацией В) Простая линейная регрессия С) Комбинированная регуляризация D) Линейная регрессия с L1 регуляризацией

23. Напишите название метода для разделения данных на обучающую и тестовую выборки: _____

24. Выберите все этапы подготовки данных для машинного обучения: а) Нормализация б) Удаление пропусков с) Кодирование категориальных признаков d) Масштабирование

25. Верно ли утверждение: "Полиномиальная регрессия может моделировать только линейные зависимости"?

26. Какой метод используется для оценки качества регрессионной модели? а) R-squared б) Accuracy с) Precision d) Recall

27. Установите соответствие между терминами и их определениями:

1. Признак
2. Целевая переменная
3. Выборка
4. Модель

А) Набор данных для обучения В) Характеристика объекта С) Прогнозируемое значение D) Математическое представление зависимостей

28. Напишите название метрики для оценки качества регрессии: _____

29. Выберите все методы регуляризации: а) L1 регуляризация б) L2 регуляризация с) Эластичная сеть d) Dropout

30. Верно ли утверждение: "Feature Engineering всегда улучшает качество модели"?
31. Какая функция используется для стандартизации данных? а) StandardScaler б) MinMaxScaler в) Normalizer д) RobustScaler
32. Установите соответствие между типами признаков и методами их обработки:
1. Числовые
 2. Категориальные
 3. Временные
 4. Текстовые
- А) OneHotEncoder В) StandardScaler С) DateTimeEncoder D) TfidfVectorizer
33. Напишите название метода для поиска оптимальных гиперпараметров:
-
34. Выберите все методы обработки пропущенных значений: а) Удаление б) Заполнение средним в) Заполнение медианой д) Интерполяция
35. Верно ли утверждение: "Cross-validation помогает избежать переобучения модели"?
36. Какой метод используется для оценки качества классификационной модели? а) ROC-AUC б) R-squared в) MSE д) RMSE
37. Установите соответствие между методами классификации и их характеристиками:
1. Логистическая регрессия
 2. К-ближайших соседей
 3. Случайный лес
 4. SVM
- А) Непараметрический метод В) Линейный классификатор С) Ансамблевый метод D) Метод опорных векторов
38. Напишите название метрики для бинарной классификации: _____
39. Выберите все параметры метода k-ближайших соседей: а) Число соседей б) Метрика расстояния в) Веса соседей д) Алгоритм поиска
40. Верно ли утверждение: "Логистическая регрессия может использоваться только для бинарной классификации"?
41. Какое распределение чаще всего встречается в природе? а) Нормальное б) Биномиальное в) Пуассона д) Экспоненциальное
42. Установите соответствие между статистическими тестами и их применением:
1. t-тест
 2. хи-квадрат
 3. ANOVA
 4. тест Шапиро-Уилка
- А) Проверка нормальности В) Сравнение средних С) Анализ категориальных данных D) Сравнение нескольких групп
43. Напишите название параметра нормального распределения, характеризующего разброс: _____
44. Выберите все свойства нормального распределения: а) Симметричность б) Унимодальность в) Непрерывность д) Определённость на всей числовой оси
45. Верно ли утверждение: "Центральная предельная теорема работает только для больших выборок"?
46. Какой метод используется для визуализации распределения данных? а) Гистограмма б) Scatter plot в) Line plot д) Bar plot
47. Установите соответствие между параметрами и их статистическими характеристиками:
1. Среднее
 2. Медиана

3. Мода
4. Дисперсия
- A) Мера разброса B) Центр распределения C) Наиболее частое значение D) Центральное значение
48. Напишите название графика для визуализации выбросов: _____
49. Выберите все методы проверки гипотез: а) Параметрические б) Непараметрические с) Байесовские d) Ранговые
50. Верно ли утверждение: "Доверительный интервал 95% означает, что истинное значение параметра всегда находится внутри интервала"?
51. Какой объект в Pandas используется для работы с одномерными данными? а) Series б) DataFrame с) Array d) List
52. Установите соответствие между операциями Pandas и их назначением:
 1. merge
 2. concat
 3. groupby
 4. pivot
- A) Группировка данных B) Объединение по ключу C) Преобразование формата таблицы D) Объединение по индексу
53. Напишите название метода для удаления дубликатов в DataFrame: _____
54. Выберите все типы объединения таблиц в Pandas: а) inner б) outer с) left d) right
55. Верно ли утверждение: "При использовании метода groupby() всегда необходимо указывать агрегирующую функцию"?
56. Какой метод используется для фильтрации строк в DataFrame? а) loc б) iloc с) filter d) select
57. Установите соответствие между функциями агрегации и их результатами:
 1. sum()
 2. mean()
 3. count()
 4. max()
- A) Максимальное значение B) Сумма значений C) Количество значений D) Среднее значение
58. Напишите название метода для заполнения пропущенных значений: _____
59. Выберите все способы индексации в Pandas: а) По метке б) По позиции с) По условию d) По срезу
60. Верно ли утверждение: "DataFrame может содержать данные разных типов в одном столбце"?
61. Какой тип графика лучше использовать для визуализации временных рядов? а) Line plot б) Scatter plot с) Bar plot d) Pie chart
62. Установите соответствие между методами анализа данных и их применением:
 1. describe()
 2. info()
 3. value_counts()
 4. corr()
- A) Подсчет уникальных значений B) Статистические характеристики C) Информация о DataFrame D) Корреляционная матрица
63. Напишите название метода для изменения типа данных столбца: _____
64. Выберите все операции для работы с временными данными в Pandas: а) Преобразование формата б) Извлечение компонент даты с) Сдвиг временного ряда d) Ресемплинг
65. Верно ли утверждение: "Метод apply() может применяться только к отдельным столбцам DataFrame"?

66. Какая функция используется для чтения CSV файла в Pandas? a) read_csv() b) load_csv() c) import_csv() d) open_csv()
67. Установите соответствие между типами данных Pandas и их Python-аналогами:
1. object
 2. int64
 3. float64
 4. datetime64
- A) Целые числа B) Строки C) Дата и время D) Числа с плавающей точкой
68. Напишите название метода для сортировки DataFrame: _____
69. Выберите все способы обработки категориальных данных: a) Label Encoding b) One-Hot Encoding c) Ordinal Encoding d) Binary Encoding
70. Верно ли утверждение: "Метод drop() всегда необратимо удаляет данные из DataFrame"?
71. Какая библиотека используется для работы с большими массивами данных в Python? a) NumPy b) SciPy c) Pandas d) Math
72. Установите соответствие между методами обработки больших данных и их назначением:
1. chunks
 2. dask
 3. pyspark
 4. vaex
- A) Распределенные вычисления B) Поточковая обработка C) Обработка по частям D) Ленивые вычисления
73. Напишите название формата для эффективного хранения табличных данных: _____
74. Выберите все характеристики больших данных: a) Объем b) Скорость c) Разнообразие d) Достоверность
75. Верно ли утверждение: "При работе с большими данными всегда нужно загружать весь датасет в память"?
76. Какой метод используется для оптимизации памяти в Pandas? a) reduce_memory_usage b) optimize_dtypes c) memory_optimize d) downcast
77. Установите соответствие между типами данных и их оптимальным представлением:
1. Целые числа
 2. Категории
 3. Даты
 4. Строки
- A) category B) int32 C) datetime64 D) object
78. Напишите название метода для параллельной обработки данных: _____
79. Выберите все методы оптимизации работы с большими данными: a) Индексация b) Партиционирование c) Кэширование d) Агрегация
80. Верно ли утверждение: "Использование категориальных типов данных всегда уменьшает потребление памяти"?
81. Какая функция используется для профилирования кода Python? a) cProfile b) timeit c) profile d) benchmark
82. Установите соответствие между форматами данных и их особенностями:
1. CSV
 2. Parquet
 3. HDF5
 4. JSON

- А) Колоночное хранение В) Текстовый формат С) Иерархическое хранение D) Структурированный текст
83. Напишите название библиотеки для распределенных вычислений: _____
84. Выберите все методы оценки производительности кода: а) Измерение времени выполнения б) Профилирование памяти с) Подсчет операций d) Мониторинг CPU
85. Верно ли утверждение: "Векторизованные операции всегда быстрее циклов в Python"?
86. Какой метод используется для эффективной конкатенации строк? а) join() б) concat() с) + d) append()
87. Установите соответствие между операциями и их вычислительной сложностью:
1. Поиск в списке
 2. Поиск в словаре
 3. Сортировка
 4. Индексация
- А) $O(1)$ В) $O(n)$ С) $O(n \log n)$ D) $O(1)$
88. Напишите название метода для группировки данных по временным интервалам: _____
89. Выберите все способы оптимизации pandas DataFrame: а) Выбор оптимальных типов данных б) Удаление неиспользуемых столбцов с) Использование категориальных данных d) Индексация часто используемых столбцов
90. Верно ли утверждение: "Использование multiprocessing всегда ускоряет обработку данных"?
91. Какой метод используется для обработки потоковых данных? а) Generator б) Iterator с) List d) Array
92. Установите соответствие между библиотеками и их функциональностью:
1. NumPy
 2. Pandas
 3. Dask
 4. PySpark
- А) Распределенные вычисления В) Векторные операции С) Обработка таблиц D) Параллельные вычисления
93. Напишите название метода для чтения данных по частям: _____
94. Выберите все методы работы с временными рядами: а) Rolling б) Expanding с) Resampling d) Shifting
95. Верно ли утверждение: "Асинхронное программирование подходит для CPU-bound задач"?
96. Какой тип индекса лучше использовать для временных рядов? а) DatetimeIndex б) RangeIndex с) Int64Index d) StringIndex
97. Установите соответствие между методами агрегации и их применением:
1. rolling
 2. cumsum
 3. agg
 4. transform
- А) Скользящие окна В) Накопительная сумма С) Множественная агрегация D) Преобразование групп
98. Напишите название метода для параллельной агрегации данных: _____
99. Выберите все способы оптимизации памяти: а) Использование генераторов б) Удаление промежуточных результатов с) Применение sparse структур d) Сжатие данных
100. Верно ли утверждение: "Индексация в Pandas всегда ускоряет доступ к данным"?

101. Какой метод используется для обработки текстовых данных в больших объемах? а) map b) apply c) applymap d) transform

102. Установите соответствие между форматами хранения и их преимуществами:
1) SQL 2) NoSQL 3) Parquet 4) CSV

А) Структурированность

В) Гибкость схемы

С) Компрессия

Д) Простота

103. Напишите название метода для эффективного объединения файлов: _____

104. Выберите все методы работы с пропущенными значениями в больших данных: а) Заполнение медианой б) Интерполяция с) Предсказание д) Удаление

105. Верно ли утверждение: "Для больших данных всегда лучше использовать распределенные вычисления"?

Краткие методические указания

При подготовке к тесту необходимо обратить на содержание основных положений и выводов, объяснение явлений и фактов, уяснение практического приложения рассматриваемых теоретических вопросов по теме. В процессе этой работы студент должен стремиться понять и запомнить основные положения рассматриваемого материала, примеры, поясняющие его.

Шкала оценки

Шкала оценки

Оценка	Баллы	Описание
5	40–50	выставляется студенту, если он выполнил без существенных ошибок все задания и ответил на все поставленные вопросы, подтверждая знание материала, умение использовать нормативные документы для подтверждения правильности собственной позиции;
4	25–39	выставляется студенту, если выполнил без существенных ошибок больше половины заданий и ответил на большинство поставленных вопросы, четко представлял свою позицию, подтверждая знание материала, умение использовать нормативные документы для подтверждения правильности собственной позиции;
3	16–24	выставляется студенту, если выполнил без существенных ошибок меньше половины заданий, ответил на некоторые поставленные вопросы, подтверждая знание материала, умение использовать нормативные документы для подтверждения правильности собственной позиции;
2	7–15	выставляется студенту, если он допустил ошибки при ответах на все из поставленных в задании вопросов.