МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ (МИНОБРНАУКИ РОССИИ)

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ «ДОНБАССКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ» (ФГБОУ ВО «ДонГТУ»)

информационных технологий и автоматизации

Факультет

	производственных процессов		
Кафедра	информационных технологий		
	УТВЕРЖДАЮ и.о. проректора по учебной работе Д.В. Мулов		
PA	АБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ		
	Анализ данных		
	(наименование дисциплины)		
(2.03.01 Математика и компьютерные науки		
	(код, наименование направления)		
	Цифровые технологии в бизнесе		
	(профиль подготовки)		
	(код, наименование направления/специальности)		
	(код, наименование направления/ <u>специальности</u>)		

Квалификация	бакалавр		
	(бакалавр/специалист/магистр)		
Форма обучения	очная		
1	(owing owing proving)		

1 Цели и задачи изучения дисциплины

Цели дисциплины. Основной целью изучения дисциплины «Анализ данных» является формирование у студентов профессиональных навыков для успешного освоения современных методов и инструментов анализа данных.

Задачи изучения дисциплины:

- освоение принципов и методик анализа данных, принципов сбора и подготовки исходных данных;
 - освоение инструментов статистического анализа данных;
 - освоение методов анализа временных рядов и прогнозирования;
- освоение основных методов и алгоритмов интеллектуального анализа данных.

Дисциплина направлена на формирование универсальных (УК-1); общепрофессиональных компетенций (ОПК-1) выпускника.

2 Место дисциплины в структуре ОПОП ВО

Логико-структурный анализ дисциплины — входит в *обязательную часть Блока 1* подготовки студентов по направлению подготовки 02.03.01 Математика и компьютерные науки (профиль «Цифровые технологии в бизнесе»).

Дисциплина реализуется кафедрой информационных технологий.

Основывается на базе дисциплин: «Численные методы», «Дискретная математика», «Информатика».

Является основой для изучения следующих дисциплин: «Основы искусственного интеллекта», «Математический инструментарий и модели оценки бизнеса».

Для изучения дисциплины необходимы компетенции, сформированные у студента для решения профессиональных задач деятельности, связанных с использованием статистических методов и информационных технологий.

Курс является фундаментом для формирования у студентов навыков по использованию в практической деятельности методов анализа данных.

Общая трудоемкость освоения дисциплины составляет 4 зачетные единицы, 144 ак.ч. Программой дисциплины предусмотрены лекционные (36 ак.ч.), практические (36 ак.ч.) занятия и самостоятельная работа студента (72 ак.ч.).

Дисциплина изучается на 3 курсе в 5-м семестре.

Форма промежуточной аттестации — экзамен.

Дисциплина направлена на формирование универсальных (УК-1); общепрофессиональных компетенций (ОПК-1) выпускника.

3 Перечень результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения ОПОП ВО

Процесс изучения дисциплины «Анализ данных» направлен на формирование компетенций, представленных в таблице 1.

Таблица 1 — Компетенции, обязательные к освоению

Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач	Код компе- тенции УК-1	Код и наименование индикатора достижения компетенции УК-1.1. Знает принципы сбора, отбора и обобщения информации УК-1.2. Умеет соотносить разнородные явления и систематизировать их в рамках избранных видов профессиональной деятельности УК-1.3. Имеет практический опыт работы с информационными источниками, опыт научного поиска, создания научных текстов
Способен консультировать и использовать фундаментальные знания в области математического анализа, комплексного и функционального анализа алгебры, аналитической геометрии, дифференциальной геометрии и топологии, дифференциальных уравнений, дискретной математики и математической логики, теории вероятностей, математической статистики и случайных процессов, численных методов, теоретической механики в профессиональной деятельности	ОПК-1	ОПК-1.1. Обладает базовыми знаниями, полученными в области математических и (или) естественных наук. ОПК-1.2. Умеет использовать их в профессиональной деятельности. ОПК-1.3. Имеет навыки выбора методов решения задач профессиональной деятельности на основе теоретических знаний

4 Объём и виды занятий по дисциплине

Общая трудоёмкость учебной дисциплины составляет 4 зачётные единицы, 144 ак.ч.

Самостоятельная работа студента (СРС) включает проработку материалов лекций, подготовку к практическим занятиям, устному опросу, текущему контролю, самостоятельное изучение материала и подготовку к экзамену.

При организации внеаудиторной самостоятельной работы по данной дисциплине используются формы и распределение бюджета времени на СРС для очной формы обучения в соответствии с таблицей 2.

Таблица 2 — Распределение бюджета времени на СРС

Вид учебной работы	Всего ак.ч.	Ак.ч. по семестрам 5
Аудиторная работа, в том числе:	72	72
Лекции (Л)	36	36
Практические занятия (ПЗ)	36	36
Лабораторные работы (ЛР)	-	-
Курсовая работа/курсовой проект	-	-
Самостоятельная работа студентов (СРС), в том числе:	72	72
Подготовка к лекциям	9	9
Подготовка к лабораторным работам	-	-
Подготовка к практическим занятиям / семинарам	36	36
Выполнение курсовой работы / проекта	-	-
Расчетно-графическая работа (РГР)	-	-
Реферат (индивидуальное задание)	-	-
Домашнее задание	-	-
Подготовка к контрольной работе	-	-
Подготовка к коллоквиуму	-	-
Аналитический информационный поиск	12	12
Работа в библиотеке	8	8
Подготовка к экзамену	7	7
Промежуточная аттестация – экзамен (Э)	Э	Э
Общая трудоемкость дисциплины		
ак.ч.	144	144
3.e.	4	4

5 Содержание дисциплины

С целью освоения компетенций, приведенных в п.3, дисциплина разбита на 6 тем:

- тема 1 (Первичные методы статистической обработки данных);
- тема 2 (Корреляционный анализ);
- тема 3 (Регрессионный анализ);
- тема 4 (Дисперсионный анализ);
- тема 5 (Методы классификации);
- тема 6 (Кластерный анализ).

Виды занятий по дисциплине и распределение аудиторных часов для очной формы приведены в таблице 3.

Таблица 3 — Виды занятий по дисциплине и распределение аудиторных часов (очная форма обучения)

№ п/п	Наименование темы (раздела) дисциплины	Содержание лекционных занятий	Трудоемкость в ак.ч.	Темы практических за- нятий	Трудоемкость в ак.ч.	Тема лабораторных занятий	Трудоемкость в ак.ч.
	Первичные ме- тоды статисти-	Предмет и задачи дисциплины. Расчет параметров описательной статистики. Предварительный анализ данных. Обнаружение влиятельных		Определение набора описательных характеристик	2		
1	ческой обработ-ки данных	наблюдений и выбросов. Определение показателей, характеризующих тенденцию динамики	6	Определение показателей, характеризующих тенденцию динамики	4	-	_
2	Корреляцион- ный анализ	Корреляционная и ковариационная матрицы. Коэффициент детерминации. Ранговый коэффициент корреляции Спирмена, коэффициент ранговой корреляции Кендалла	4	Корреляционный анализ	4	-	_
3	Регрессионный анализ	Линейная регрессия. Анализ случайной составляющей уравнения регрессии. Нелинейные регрессионные модели.	6	Регрессионный анализ	6		
4	Дисперсионный анализ	Дисперсионный анализ. Однофакторный дисперсионный анализ. Множественный дисперсионный анализ.	6	Дисперсионный анализ	6		
7	Методы класси- фикации	Понятие о классификации данных и основные алгоритмы классификации. Логистическая регрессия, наивный байесовский классификатор, k-ближайшие соседи, деревья принятия решений, случайный лес, искусственные нейронные сети, машина опорных векторов, стохастический градиентный спуск	6	Классификация мето- дом случайного леса	6		

7

№ п/п	Наименование темы (раздела) дисциплины	Содержание лекционных занятий	Трудоемкость в ак.ч.	Темы практических за- нятий	Трудоемкость в ак.ч.	Тема лабораторных занятий	Трудоемкость в ак.ч.
6	Кластерный ана- лиз	Понятие о кластеризации. Методы кластеризации. Кластеризация иерархическим методом и методом k-средних.	4	V поставилогия навовую		I	_
7	OLAP-технология	Понятие оперативной аналитической обработки данных. Компоненты OLAP-системы. Варианты хранения данных в OLAP-системах. Хранилища и озера данных	2	Кластеризация иерархическим методом и методом к-средних.	6		
	Всего аудиторных часов		36	36		-	

6 Фонд оценочных средств для проведения текущего контроля успеваемости и промежуточной аттестации по итогам освоения дисциплины

6.1 Критерии оценивания

В соответствии с Положением о кредитно-модульной системе организации образовательного процесса ФГБОУ ВО «ДонГТУ» (https://www.dstu.education/images/structure/license_certificate/polog_kred_modu l.pdf) при оценивании сформированности компетенций по дисциплине используется 100-балльная шкала.

Перечень работ по дисциплине и способы оценивания знаний приведены в таблице 4.

Таблица 4 — Перечень работ по дисциплине и способы оценивания знаний

Вид учебной работы	Способ оценивания	Количество баллов
Выполнение практических работ	Предоставление отче- тов	50–80
Выполнение тестового контроля или устного опроса	Более 50% правильных ответов	10–20
Итого	-	60–100

Экзамен проставляется автоматически, если студент набрал в течение семестра не менее 60 баллов и отчитался за каждую контрольную точку. Минимальное количество баллов по каждому из видов текущей работы составляет 60 % от максимального.

Экзамен по дисциплине «Анализ данных» проводится по результатам работы в семестре. В случае, если полученная в семестре сумма баллов не устраивает студента, во время экзамена студент имеет право повысить итоговую оценку в форме устного экзамена по приведенным ниже вопросам (п.п. 6.3).

Шкала оценивания знаний при проведении промежуточной аттестации приведена в таблице 5.

Таблица 5 — Шкала оценивания знаний

Сумма баллов за все виды учебной	Оценка по национальной шкале
деятельности	зачёт/экзамен
0–59	Не зачтено/неудовлетворительно
60–73	Зачтено/удовлетворительно
74–89	Зачтено/хорошо
90–100	Зачтено/отлично

6.2 Оценочные средства для самостоятельной работы и текущего контроля успеваемости: тестовый контроль

п/п 2 3 Тема 1 Первичные методы статистической обработки данных 1 Какой инструмент электронных таблиц позволяет определить статистические характеристики ряда данных? а) Описательная статистика; 2 Каково назначение инструмента электронных таблиц Описательная статистика? а) построение уравнения регрессии, наилучинобразом описывающего анализируемые данных статистических харанных статистических харанных для диапазона данных; 8) заполнение диапазона данных; в) заполнение диапазона данных случайными числами, извлеченными из одного или неском ких распределений 3 Что такое временной ряд? а) совокупность значений нескольких незави мых факторов в один момент времени;	ые;
Тема 1 Первичные методы статистической обработки данных 1 Какой инструмент электронных таблиц позволяет определить статистические характеристики ряда данных? 2 Каково назначение инструмента электронных таблиц Описательная статистика? 3 Описательная статистика; 6) Корреляция; 8) Ковариация; 7) Регрессия а) построение уравнения регрессии, наилучинобразом описывающего анализируемые данных статистических харатеристик для диапазона данных; 8) определение основных статистических харатеристик для диапазона данных; 8) заполнение диапазона данных случайными числами, извлеченными из одного или неском ких распределений 3 Что такое временной ряд?	ые;
1 Какой инструмент электрон- ных таблиц позволяет опреде- лить статистические характе- ристики ряда данных? 2 Каково назначение инструмен- та электронных таблиц Описа- тельная статистика? 5 Описательная статистика; 6 Корреляция; в) Ковариация; г) Регрессия а) построение уравнения регрессии, наилучинобразом описывающего анализируемые данны б) определение основных статистических хара теристик для диапазона данных; в) заполнение диапазона данных случайными числами, извлеченными из одного или неском ких распределений 3 Что такое временной ряд? а) Описательная статистика; б) Корреляция; в) Ковариация; г) Регрессия а) построение уравнения регрессии, наилучинобразом описывающего анализируемые данны б) определение основных статистических хара теристик для диапазона данных случайными числами, извлеченными из одного или неском ких распределений 3 Что такое временной ряд?	ые;
ных таблиц позволяет определить статистические характеристики ряда данных? 2 Каково назначение инструмента электронных таблиц Описательная статистика? 3 Что такое временной ряд? (б) Корреляция; в) Ковариация; г) Регрессия а) построение уравнения регрессии, наилучино образом описывающего анализируемые данных образом описывающего анализируемые данных статистических хара теристик для диапазона данных; в) заполнение диапазона данных; в) заполнение диапазона данных случайными числами, извлеченными из одного или неском ких распределений а) совокупность значений нескольких незави	ые;
лить статистические характеристики ряда данных? 2 Каково назначение инструмента электронных таблиц Описательная статистика? 3 Определение основных статистических хара теристик для диапазона данных; в заполнение диапазона данных случайными числами, извлеченными из одного или неском ких распределений 3 Что такое временной ряд? В) Ковариация; г) Регрессия а) построение уравнения регрессии, наилучите образом описывающего анализируемые данных статистических хара теристик для диапазона данных; в заполнение диапазона данных случайными числами, извлеченными из одного или неском ких распределений 3 Что такое временной ряд?	ые;
ристики ряда данных? 2 Каково назначение инструмента электронных таблиц Описательная статистика? 3 Определение основных статистических хар теристик для диапазона данных; в) заполнение диапазона данных случайными числами, извлеченными из одного или неском ких распределений 3 Что такое временной ряд? 1 Регрессия а) Построение уравнения регрессии, наилучино образом описывающего анализируемые данн б) определение основных статистических хар теристик для диапазона данных; в) заполнение диапазона данных случайными числами, извлеченными из одного или неском ких распределений 3 Совокупность значений нескольких незави	ые;
2 Каково назначение инструмента электронных таблиц Описательная статистика? а) построение уравнения регрессии, наилучитобразом описывающего анализируемые данный б) определение основных статистических хартеристик для диапазона данных; в) заполнение диапазона данных случайными числами, извлеченными из одного или неском ких распределений 3 Что такое временной ряд? а) совокупность значений нескольких незави	ые;
та электронных таблиц Описательная статистика? б) определение основных статистических хартеристик для диапазона данных; в) заполнение диапазона данных случайными числами, извлеченными из одного или неском ких распределений 3 Что такое временной ряд? а) совокупность значений нескольких незави	ые;
теристик для диапазона данных; в) заполнение диапазона данных случайными числами, извлеченными из одного или неском ких распределений 3 Что такое временной ряд? а) совокупность значений нескольких незави	оак-
в) заполнение диапазона данных случайными числами, извлеченными из одного или неском ких распределений 3 Что такое временной ряд? а) совокупность значений нескольких незави	
числами, извлеченными из одного или неском ких распределений 3 Что такое временной ряд? а) совокупность значений нескольких незави	
ких распределений 3 Что такое временной ряд? а) совокупность значений нескольких незави	1
3 Что такое временной ряд? а) совокупность значений нескольких незави	пь-
мых факторов в один момент времени;	си-
1 1 1	
б) совокупность значений нескольких незави	си-
мых факторов за несколько последовательны	\mathbf{X}
моментов времени;	
в) совокупность значений какого-либо показа	ателя
за несколько последовательных моментов ил	И
периодов времени	
4 Что показывает абсолютный а) прирост уровня ряда относительно предыд	уще-
цепной прирост? го периода времени;	_
б) уровень изменения величины относительн	o 6a-
зового периода времени;	
в) ускорение или замедление тенденции изуч	ae-
мого процесса	
5 Что показывает цепной коэф- а) прирост показателя в анализируемый по	-
фициент прироста? времени относительно базового уровня,	зыра-
женный в процентах;	17/11/10
б) прирост уровня ряда относительно предыд	јуще-
го значения, выраженный в процентах; в) увеличение уровня ряда относительно процентах	oenti-
дущего значения, выраженное в процентах;	лоды-
г) рост показателя в анализируемый период в	ne-
мени относительно базового уровня, выраже	
в процентах	.11111111
6 Что не относится к мерам цен- а) размах;	
тральной тенденции? б) среднее арифметическое;	
в) медиана;	
г) усеченное среднее;	
д) мода	

No	_	_
п/п	Содержание вопроса	Варианты ответа
1	2	3
7	Что не относится к мерам рассеяния?	а) размах; б) дисперсия; в) стандартное отклонение; г) медианное абсолютное отклонение; д) усеченное среднее;
		е) межквартильный размах
8	Какую из перечисленного набора описательных характеристик не вычисляет функция языка R summary()?	a) Мин.; б) 1st Qu; в) Медиана; г) 3-й Qu; д) Макс.; е) размах
9	Каково назначение функции describe() языка R?	а) получение статистических характеристик набора данных; б) удаление неиспользуемых уровней из факторов во фрейме данных; в) определение усеченного среднего
10	Что такое выбросы при обработке данных?	а) отсутствующие значения по одному или нескольким элементам; б) экстремальные значения во входных данных, которые находятся далеко за пределами других наблюдений; в) факторы, которые следует исключить из рассмотрения ввиду наличия мультиколлинеарности
11	Что такое влиятельное наблюдение при обработке данных?	а) наблюдение для статистического вычисления, удаление которого из набора данных заметно изменяет результат вычисления; б) наблюдение, в котором отсутствуют значения по одному или нескольким элементам; в) наблюдение за фактором, оказывающим наибольшее влияние на анализируемый результирующий показатель
12	Что такое очистка данных?	а) процесс удаления из базы данных информации, срок хранения которой истек; б) процесс выявления и исправления ошибок, несоответствий данных с целью улучшения их качества, иногда классифицируется как составная часть интеллектуального анализа данных; в) экстремальные значения во входных данных, которые находятся далеко за пределами других наблюдений

№ п/п	Содержание вопроса	Варианты ответа
1	2	3
	Тема 2 К	орреляционный анализ
1	Каково назначение корреляционного анализа?	а) вычисление среднего произведения отклонений точек данных от относительных средних; б) количественная оценка взаимосвязи двух наборов данных; в) проверка гипотезы о сходстве средних значений двух или более выборок; г) подбор графика для набора наблюдений с помощью метода наименьших квадратов
2	Что характеризует коэффици- ент корреляции?	а) величину отклонения значений, рассчитанных по уравнению тренда, от исходных данных; б) часть дисперсии показателя, объясняемой регрессией; в) тесноту линейной связи между признаками
3	В каких пределах может изменяться значение величины коэффициента корреляции?	a) от 0 до 0,5; б) от 0 до 1; в) от 0 до 2; г) от -1 до 1
	Тема 3 Н	Регрессионный анализ
1	Каково назначение регресси-	а) вычисление среднего произведения отклонений
	онного анализа?	точек данных от относительных средних; б) подбор графика для набора наблюдений с помощью метода наименьших квадратов; в) проверка гипотезы о сходстве средних значений двух или более выборок; г) количественная оценка взаимосвязи двух наборов данных
2	Что характеризует коэффици- ент детерминации?	а) тесноту линейной связи между признаками; б) величину отклонения значений, рассчитанных по уравнению тренда, от исходных данных; в) часть дисперсии показателя, объясняемой регрессией
3	Какие функции не относятся к классу нелинейных относительно включенных в анализ независимых переменных, но линейных по оцениваемым параметрам?	а) экспоненциальная; б) гиперболическая; в) полулогарифмическая; г) полиномиальная
4	Какие функции не относятся к классу регрессий, нелинейных по оцениваемым параметрам?	а) степенная; б) гиперболическая; в) показательная; г) экспоненциальная
5	Каково назначение функции языка R predict()?	а) прогнозирование значений на основе входных данных; б) получение статистических характеристик набора данных; в) определение усеченного среднего

No	Содержание вопроса	Варианты ответа
п/п		-
1	2	3
1	I	исперсионный анализ
1	Каково назначение дисперсионного анализа?	а) вычисление среднего произведения отклонений точек данных от относительных средних; б) количественная оценка взаимосвязи двух наборов данных;
		в) проверка гипотезы о сходстве средних значений двух или более выборок
2	Что положено в основу дисперсионного анализа?	а) разложение общей дисперсии зависимой переменной на дисперсию за счет разбиения на группы и дисперсию за счет остальных факторов; б) сравнение дисперсий двух или более анализируемых выборок; в) сравнение средних значений двух или более анализируемых выборок
3	Какой критерий используется для сравнения внутри- и меж-групповых дисперсий?	а) критерий Стьюдента; б) критерий хи-квадрат Пирсона; в) критерий Фишера
4	Какая функция языка R позволяет выполнить дисперсионный анализ?	a) sd(); b) lm(); b) var()
	Тема 5 М	етоды классификации
1	Как называется распределение множества разнородных объектов по группам на основании каких-то признаков?	Введите ответ
2	Какой метод используется для классификации, когда переменная ответа имеет категориальные значения?	а) логистическая регрессия; б) наивный байесовский классификатор; в) метод ближайших соседей; г) деревья решений
3	Какая функция языка R позволяет выполнить классификацию методом случайного леса?	a) rpart(); б) randomForest(); в) multinom()
		Кластерный анализ
1	Как называется группа объектов, имеющих общие свойства?	Введите ответ
2	Как называется статистический метод исследований, позволяющий объединять или группировать набор объектов в небольшие группы, которые отличаются по характеристикам от других таких же различных групп?	Введите ответ

№		
п/п	Содержание вопроса	Варианты ответа
1	2	3
3	Что такое метод кластеризации?	а) способ вычисления расстояний между кластерами;б) способ последовательного слияния кластеров;в) способ последовательного разделения кластеров
4	Каковы наиболее популярные методы кластеризации?	а) иерархическая кластеризация и метод ближай- ших соседей; б) иерархическая кластеризация и метод k- средних (k-means); в) иерархическая кластеризация и метод опорных векторов
5	Что такое кластеризация иерархическим методом?	а) общее семейство алгоритмов кластеризации, основанное на минимизации суммарных квадратичных отклонений точек кластеров от средних координат этих кластеров; б) общее семейство алгоритмов кластеризации, основанное на определении областей с высокой плотностью; в) общее семейство алгоритмов кластеризации, которые создают вложенные кластеры путем их последовательного слияния или разделения
6	Как называется визуализатор, используемый в языке R для представления результатов иерархической кластеризации?	Введите ответ
7	Как называется метод, основанный на минимизации суммарных квадратичных отклонений точек кластеров от центроидов (средних координат) этих кластеров?	Введите ответ
	Тема 7	7 OLAP-технология
1	Каково назначение техноло- гии OLAP?	а) подготовка суммарной (агрегированной) информации на основе больших массивов данных, структурированных по многомерному принципу; б) организация обработки в реляционных базах данных в) орга;низация обработки в нереляционных базах данных
2	Какова основная причина использования технологии OLAP для обработки запросов?	а) возможность организовать обработку в нереляционных базах данных; б) скорость; в) удобство интерфейса

3	Что такое хранилище данных?	а) база данных, предназначенная для хранения
		резервных копий документов организации;
		б) база данных, специально разработанная и
		предназначенная для длительного хранения ар-
		хивных документов организации со сроком хра-
		нения не менее двух лет;
		в) предметно-ориентированная информационная
		база данных, специально разработанная и предна-
		значенная для подготовки отчётов и бизнес-ана-
		лиза с целью поддержки принятия решений в ор-
		ганизации
4	Что такое озеро данных?	а) централизованное хранилище, позволяющее
		хранить большие объемы структурированных
		данных;
		б) централизованное хранилище, позволяющее
		хранить большие объемы неструктурированных
		данных;
		в) централизованное хранилище, которое позво-
		ляет хранить все структурированные и неструк-
		турированные данные в любом масштабе
5	В чем заключается основной	а) необработанные данные хранятся без контроля
	недостаток озер данных?	за их содержимым;
		б) озера данных невозможно развертывать в об-
		лаке;
		в) не позволяет хранить нереляционные данные
		из мобильных приложений, социальных сетей и
		устройств ІоТ

6.3 Вопросы для подготовки к экзамену

- 1) Что относится к мерам центральной тенденции?
- 2) Что относится к мерам рассеяния?
- 3) Какой инструмент электронных таблиц позволяет определить статистические характеристики ряда данных?
- 4) Какие описательные характеристики вычисляют функции языка R summary() и describe()?
 - 5) Что такое временной ряд?
 - 6) Что показывает абсолютный цепной прирост?
 - 7) Что показывает цепной коэффициент прироста?
- 8) Что такое выбросы и влиятельные наблюдения при обработке данных?
 - 9) Что такое очистка данных?
 - 10) Каково назначение корреляционного анализа?
 - 11) Что характеризует коэффициент корреляции?
- 12) В каких пределах может изменяться значение величины коэффициента корреляции?
 - 13) Каково назначение функции predict()?
 - 14) Каково назначение регрессионного анализа?
 - 15) Что характеризует коэффициент детерминации?
- 16) На какие классы подразделяются нелинейные регрессионные модели?
- 17) Каково назначение дисперсионного анализа и что положено в его основу?
- 18) Какой критерий используется для сравнения внутри- и межгрупповых дисперсий?
- 19) Какая функция языка R позволяет выполнить дисперсионный анализ?
 - 20) Что такое кластеризация?
 - 21) Что такое метод кластеризации?
 - 22) Каковы наиболее популярные методы кластеризации?
 - 23) Что такое кластеризация иерархическим методом?
 - 24) Каковы основные утверждения о дендрограмме?
 - 25) На чем основан метод k-средних?
 - 26) Что такое классификация?
- 27) Какой метод классификации целесообразно использовать, если набор данных содержит выбросы?
 - 28) Каковы основные утверждения об алгоритме случайного леса?
- 29) Какая функция используется в R для создания и анализа случайного леса?
- 30) Какой метод используется для классификации, когда переменная ответа имеет категориальные значения?
 - 31) Каково назначение технологии OLAP?

- 32) Какова основная причина использования технологии OLAP для обработки запросов?
 - 33) Что такое хранилище данных?
 - 34) Что такое озеро данных?
 - 35) В чем заключается основной недостаток озер данных?

6.4 Примерная тематика курсовых работ

Курсовые работы не предусмотрены.

7 Учебно-методическое и информационное обеспечение дисциплины

7.1 Рекомендуемая литература

Основная литература Основная литература

- 1. Гришин, В. А. Методы обработки данных и моделирование на языке R: учеб. методическое пособие / В. А. Гришин, М. С. Тихов. Нижний Новгород, 2019. 54 с. URL: http://old.lib.unn.ru/students/src/grishin_tihov_R.pdf. Текст: электронный.
- 2. Поздняков, И. Статистика, R и анализ данных / И. Поздняков, T. Петухов, 2020. URL : https://pozdniakov.github.io/stats/. Текст : электронный.
- 3. Как провести анализ данных в R, 2023. URL : https://sky.pro/media/kak-provesti-analiz-dannyh-v-r-2/. Текст : электронный.

Дополнительная литература

- 1. Задорожный, С. С. Статистическая обработка данных на языке R: учеб.-методическое / С. С. Задорожный.— M: Физический факультет МГУ им. М.В. Ломоносова, 2023.— 104 с. URL: https://cmp.phys.msu.ru/sites/default/files/R%20Учебное%20пособие_с_титульными стр.pdf. Текст: электронный.
- 2. Статистика и анализ данных на языке R. URL : https://rstats-at-bio-msu.netlify.app/data/2022_spring/lesson_09/2022-04-11_slides.html#1. Текст : электронный.
- 3. Язык программирования R: что делает его таким важным для анализа данных. URL: https://practicum.yandex.ru/blog/chto-takoe-yazyk-r/. Текст: электронный.

7.2 Базы данных, электронно-библиотечные системы, информационно-справочные и поисковые системы

- 1. Научная библиотека ДонГТУ: официальный сайт. Алчевск. URL: library.dstu.education. Текст: электронный.
- 2. Научно-техническая библиотека БГТУ им. Шухова : официальный сайт. Белгород. URL: http://ntb.bstu.ru/jirbis2/. Текст : электронный.
- 3. Консультант студента : электронно-библиотечная система. Mockba. URL: http://www.studentlibrary.ru/cgi-bin/mb4x. Текст : электронный.
- 4. Университетская библиотека онлайн : электронно-библиотечная система. URL: http://biblioclub.ru/index.php?page=main_ub_red. Текст : электронный.
- 5. IPR BOOKS : электронно-библиотечная система. Красногорск. URL: http://www.iprbookshop.ru/. Текст : электронный.

8 Материально-техническое обеспечение дисциплины

Материально-техническая база обеспечивает проведение всех видов деятельности в процессе обучения, соответствует требованиям ФГОС ВО.

Материально-техническое обеспечение представлено в таблице 8.

Таблица 8 — Материально-техническое обеспечение

Наименование оборудованных учебных кабинетов	Адрес (местопо- ложение) учеб- ных кабинетов
Специальные помещения: Компьютерный класс с мультимедийным оборудованием (25 посадочных мест), оборудованный учебной мебелью, компьютерами с неограниченным доступом к сети Интернет, включая доступ к	ауд. <u>412</u> корп. 2
ЭБС: компьютер – 14 шт., мультимедийный проектор, проекционный экран, веб-камера, колонки, микрофон, принтер Pantum P2516, доска для написания мелом	
Компьютерный класс кафедры ИТ (25 посадочных мест), оборудованный учебной мебелью, компьютерами с неограниченным доступом к сети Интернет, включая доступ к ЭБС: компьютер – 14 шт., интерактивная панель, принтер Pantum P2516	ауд. <u>314</u> корп. 2
Компьютерный класс кафедры ИТ (25 посадочных мест), оборудованный учебной мебелью, компьютерами с неограниченным доступом к сети Интернет, включая доступ к ЭБС: персональный компьютер Intel Celeron 420 / ECS 945GCT-M2 / DDR2 2GB / HDD Hitachi 120 GB / TFT Монитор Hanns.G 18.5" — 14 шт., принтер Canon LBP-810 — 1 шт., принтер Epson LX300 — 1 шт., сканер A4 HP-400 — 1 шт., мультимедийная доска — 1 шт., столы компьютерные — 27 шт.; парты — 5 шт.; стулья — 30 шт.	ауд. <u>302</u> корп. 2

Лист согласования рабочей программы дисциплины

Разработал		
к.т.н., доцент кафедры		
информационных технологий	It heers	Н.Н. Лепило
(должность)	(подпись)	(Ф.И.О.)
(должность)	(подпись)	(Ф.И.О.)
(должность)	(подпись)	(Ф.И.О.)
И.о. заведующего кафедрой информационных технологий	Matur	<u>А.Н. Баранов</u> (Ф.И.О.)
Протокол № 1 заседания кафедры	уподпись)	(Ψ.Μ.Ο.)
информационных технологий		от <u>26.08.2024г</u> .
Согласовано		
Председатель методической		
комиссии по направлению подготовки 02.03.01 Математика и компьютерные науки		
(профиль: «Цифровые технологии в бизнесе	») <u>{{ML}}</u> (подпись)	<u>Н.Н. Лепило</u> (Ф.И.О.)
Начальник учебно-методического центра	(подпись	О.А.Коваленко Ф.И.О.)

Лист изменений и дополнений

Номер изменения, дата внесения изменения, номер страницы для внесения					
изменений					
ДО ВНЕСЕНИЯ ИЗМЕНЕНИЙ:	ПОСЛЕ ВНЕСЕНИЯ ИЗМЕНЕНИЙ:				
Основание:					
Полицен лице отратетрациоте за визосние изменений					
Подпись лица, ответственного за внесение изменений					