

Министерство образования и науки Астраханской области
Государственное бюджетное образовательное учреждение
Астраханской области высшего образования
«Астраханский государственный архитектурно-строительный
университет»
(ГБОУ АО ВО «АГАСУ»)



РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Наименование дисциплины

«Системы обработки больших данных»

(указывается наименование в соответствии с учебным планом)

По направлению подготовки

09.04.02 «Информационные системы и технологии»

(указывается наименование направления подготовки в соответствии с ФГОС ВО)

Направленность (профиль)

«Искусственный интеллект в проектировании городской среды»

(указывается наименование направленности (профиля) в соответствии с ОПОП)

Кафедра Системы автоматизированного проектирования и моделирования

Квалификация выпускника *магистр*

Разработчик:

К.т.н., доцент
(занимаемая должность,
учёная степень и учёное звание)



(подпись)

/С.В. Окладникова/
(И. О. Ф.)

Рабочая программа рассмотрена и утверждена на заседании кафедры
«Системы автоматизированного проектирования и моделирования»
протокол № 9 от 22.04.2024 г.

И.о. заведующего кафедрой



(подпись)

/В.В. Соболева /
(И. О. Ф.)

Согласовано:

Председатель МКН «Информационные системы и технологии»
направленность(профиль) «Искусственный интеллект в проектировании городской среды»



(подпись)

/В.В. Соболева/
И.О.Ф.

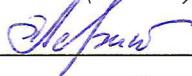
Начальник УМУ



(подпись)

| В.А. Беланов |
(И.О.Ф.)

Специалист УМУ



(подпись)

| С.А. Нарин |
(И.О.Ф.)

Начальник УИТ



(подпись)

| А.И. Узра |
(И.О.Ф.)

Заведующая научной библиотекой



(подпись)

| Л.С. Тобринова |
(И.О.Ф.)

Содержание

1. Цель освоения дисциплины	4
2. Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения образовательной программы	4
3. Место дисциплины в структуре ОПОП бакалавриата	5
4. Объем дисциплины в зачетных единицах с указанием количества академических часов, выделенных на контактную работу обучающихся с преподавателем (по типам учебных занятий) и на самостоятельную работу обучающихся	5
5. Содержание дисциплины, структурированное по разделам с указанием отведенного на них количества академических часов и типов учебных занятий	6
5.1. Разделы дисциплины и трудоемкость по типам учебных занятий и работы обучающихся (в академических часах)	6
5.1.1. Очная форма обучения	6
5.1.2. Заочная форма обучения	6
5.1.3. Очно-заочная форма обучения	7
5.2. Содержание дисциплины, структурированное по разделам	8
5.2.1. Содержание лекционных занятий	8
5.2.2. Содержание лабораторных занятий	9
5.2.3. Содержание практических занятий	9
5.2.4. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине	10
5.2.5. Темы контрольных работ	11
5.2.6. Темы курсовых проектов/ курсовых работ	11
6. Методические указания для обучающихся по освоению дисциплины	11
7. Образовательные технологии	12
8. Учебно-методическое и информационное обеспечение дисциплины	13
8.1. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины	13
8.2. Перечень необходимого лицензионного и свободно распространяемого программного обеспечения, в том числе отечественного производства, используемого при осуществлении образовательного процесса по дисциплине	14
8.3. Перечень современных профессиональных баз данных и информационных справочных систем, доступных обучающимся при освоении дисциплины	14
9. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине	14
10. Особенности организации обучения по дисциплине для инвалидов и лиц с ограниченными возможностями здоровья	15

1. Цель освоения дисциплины

Целью освоения дисциплины «Системы обработки больших данных» является формирование компетенций у обучающихся в соответствии с требованиями Федерального государственного образовательного стандарта высшего образования по направлению подготовки 09.04.02 «Информационные системы и технологии».

2. Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения образовательной программы

В результате освоения дисциплины обучающийся должен овладеть следующими компетенциями:

ПК-1ИИП - Способен исследовать применение интеллектуальных систем для различных предметных областей

ПК-1ИИП.2 – Выбирает комплексы методов и инструментальных средств искусственного интеллекта для решения задач в зависимости от особенностей предметной области

ПК-4ИИП - Способен руководить проектами по созданию комплексных систем на основе аналитики больших данных в различных отраслях со стороны заказчика

ПК-4ИИП.1 – Руководит проектами по построению комплексных систем на основе аналитики больших данных в различных отраслях со стороны заказчика

В результате освоения дисциплин, обучающийся должен овладеть следующими результатами обучения:

Знать:

- ПК-1ИИП.2 3-1. Знает методы и инструментальные средства систем искусственного интеллекта, критерии их выбора и методы комплексирования в рамках применения интегрированных гибридных интеллектуальных систем различного назначения
- ПК-4ИИП.1 3-1. Знает методологию и принципы руководства проектами по созданию, поддержке и использованию комплексных систем на основе аналитики больших данных со стороны заказчика
- ПК-4ИИП.1 3-2. Знает специфику сфер и отраслей, для которых реализуется проект по аналитике больших данных

Уметь:

- ПК-1ИИП.2 У-1. Умеет выбирать и комплексно применять методы и инструментальные средства систем искусственного интеллекта, критерии их выбора
- ПК-4ИИП.1 У-1. Умеет решать задачи по руководству коллективной проектной деятельностью для создания, поддержки и использования комплексных систем на основе аналитики больших данных со стороны заказчика
- ПК-4ИИП.1 У-2. Умеет выявлять небольшие по масштабу проекты аналитики, которые потенциально могут представлять интерес для ряда подразделений / служб или для организации в целом
- ПК-4ИИП.1 У-3. Умеет выявлять области деловой деятельности, которые потенциально могут получить отдачу от аналитики

3. Место дисциплины в структуре ОПОП бакалавриата

Дисциплина Б1.О.6 «Системы обработки больших данных» реализуется в рамках Блока 1 «Дисциплины (модули)» обязательной части.

Дисциплина базируется на основах обучения, полученных в рамках изучения дисциплин: Специальные главы математики, Технологии программирования и инструментальные средства разработки систем искусственного интеллекта, Технологии анализа данных.

4. Объем дисциплины в зачетных единицах с указанием количества академических часов, выделенных на контактную работу обучающихся с преподавателем (по типам учебных занятий) и на самостоятельную работу обучающихся

Форма обучения	Очная	Заочная
1	2	3
Трудоемкость в зачетных единицах:	4 семестр – 3 з.е.; всего – 3 з.е.	4 семестр – 3 з.е.; всего – 3 з.е.
Лекции (Л)	4 семестр – 12 часов; всего – 12 часов	4 семестр – 2 часа; всего – 2 часа
Лабораторные занятия (ЛЗ)	учебным планом <i>не предусмотрены</i>	учебным планом <i>не предусмотрены</i>
Практические занятия (ПЗ)	4 семестр – 24 часа; всего – 24 часа	4 семестр – 8 часов; всего – 8 часов
Самостоятельная работа (СР)	4 семестр – 72 часа; всего – 72 часа	4 семестр – 98 часов; всего – 98 часов
Форма текущего контроля:		
Контрольная работа	4 семестр	4 семестр
Форма промежуточной аттестации:		
Экзамены	семестр – 4	семестр – 4
Зачеты	учебным планом <i>не предусмотрены</i>	учебным планом <i>не предусмотрены</i>
Зачет с оценкой	учебным планом <i>не предусмотрены</i>	учебным планом <i>не предусмотрены</i>
Курсовая работа	учебным планом <i>не предусмотрены</i>	учебным планом <i>не предусмотрены</i>
Курсовой проект	учебным планом <i>не предусмотрены</i>	учебным планом <i>не предусмотрены</i>

5. Содержание дисциплины, структурированное по разделам с указанием отведенного на них количества академических часов и типов учебных занятий

5.1. Разделы дисциплины и трудоемкость по типам учебных занятий и работы обучающихся (в академических часах)

5.1.1. Очная форма обучения

№ п/п	Раздел дисциплины. (по семестрам)	Всего часов на раздел	Семестр	Распределение трудоемкости раздела (в часах) по типам учебных занятий и работы обучаю- щихся				Форма текущего контроля и промежуточной аттестации
				контактная			СР	
				Л	ЛЗ	ПЗ		
1	2	3	4	5	6	7	8	9
1.	Базовые структуры и понятия языка R	27	4	3	-	6	18	контрольная работа экзамен
2.	Предобработка данных в R	27	4	3	-	6	18	
3.	Визуализация в R	27	4	3	-	6	18	
4.	Статистика в R	27	4	3	-	6	18	
Итого:		108		12	-	24	72	

5.1.2. Заочная форма обучения

№ п/п	Раздел дисциплины. (по семестрам)	Всего часов на раздел	Семестр	Распределение трудоемкости раздела (в часах) по типам учебных занятий и работы обучаю- щихся				Форма текущего контроля и промежуточной аттеста- ции
				контактная			СР	
				Л	ЛЗ	ПЗ		
1	2	3	4	5	6	7	8	9
1.	Базовые структуры и понятия языка R	27	4	-	-	2	25	контрольная работа экзамен
2.	Предобработка данных в R	27	4	1	-	2	24	
3.	Визуализация в R	27	4	-	-	2	25	
4.	Статистика в R	27	4	1	-	2	24	
Итого:		108		2	-	8	98	

5.1.3. Очно-заочная форма обучения

«ОПОП не предусмотрено»

5.2. Содержание дисциплины, структурированное по разделам

5.2.1. Содержание лекционных занятий

№	Наименование раздела дисциплины	Содержание
1	2	3
1.	Базовые структуры и понятия языка R	<p>Типы данных в R. Объекты данных. Синтаксис R. Манипуляции с данными, таблицами, датой и временем. Обработка строк. Ввод и вывод данных. Функции. Выражения.</p> <p><i>Методы и инструментальные средства систем искусственного интеллекта, критерии их выбора и методы комплексирования в рамках применения интегрированных гибридных интеллектуальных систем различного назначения.</i></p> <p><i>Методология и принципы руководства проектами по созданию, поддержке и использованию комплексных систем на основе аналитики больших данных со стороны заказчика.</i></p> <p><i>Специфика сфер и отраслей, для которых реализуется проект по аналитике больших данных.</i></p>
2.	Предобработка данных в R	<p>Методы предобработки данных. Этапы предобработки данных: очистка данных, редактирование данных, заполнение пропусков, форматирование данных, отбор признаков данных.</p> <p><i>Методы и инструментальные средства систем искусственного интеллекта, критерии их выбора и методы комплексирования в рамках применения интегрированных гибридных интеллектуальных систем различного назначения.</i></p> <p><i>Методология и принципы руководства проектами по созданию, поддержке и использованию комплексных систем на основе аналитики больших данных со стороны заказчика.</i></p> <p><i>Специфика сфер и отраслей, для которых реализуется проект по аналитике больших данных.</i></p>
3.	Визуализация в R	<p>Философия визуализации в R. Инструменты визуализации в R. Ggplot2. Диаграммы, гистограммы, тепловые карты.</p> <p><i>Методы и инструментальные средства систем искусственного интеллекта, критерии их выбора и методы комплексирования в рамках применения интегрированных гибридных интеллектуальных систем различного назначения.</i></p> <p><i>Методология и принципы руководства проектами по созданию, поддержке и использованию комплексных систем на основе аналитики больших данных со стороны заказчика.</i></p> <p><i>Специфика сфер и отраслей, для которых реализуется проект по аналитике больших данных.</i></p>
4.	Статистика в R	<p>Статистические гипотезы. Статистические ошибки. Статистические характеристики. Статистическая разведка. Оценки неизвестных параметров в R. Анализ временных рядов. Проверка статистических гипотез.</p> <p><i>Методы и инструментальные средства систем искусственного интеллекта, критерии их выбора и методы комплексирования в рамках применения интегрированных гибридных интеллектуальных систем различного назначения.</i></p> <p><i>Методология и принципы руководства проектами по созданию, поддержке и использованию комплексных систем на основе аналитики больших данных со стороны заказчика.</i></p>

№	Наименование раздела дисциплины	Содержание
1	2	3
		<i>нию, поддержке и использованию комплексных систем на основе аналитики больших данных со стороны заказчика. Специфика сфер и отраслей, для которых реализуется проект по аналитике больших данных.</i>

5.2.2. Содержание лабораторных занятий

Учебным планом не предусмотрены

5.2.3. Содержание практических занятий

№	Наименование раздела дисциплины	Содержание
1	2	3
1.	Базовые структуры и понятия языка R	<i>Входное тестирование. Загрузка данных в R Выбор и комплексное применение методов и инструментальных средств систем искусственного интеллекта, критериев их выбора решение задач по руководству коллективной проектной деятельностью для создания, поддержки и использования комплексных систем на основе аналитики больших данных со стороны заказчика. Выявление небольших по масштабу проектов аналитики, которые потенциально могут представлять интерес для ряда подразделений / служб или для организации в целом Выявление областей деловой деятельности, которые потенциально могут получить отдачу от аналитики.</i>
2.	Предобработка данных в R	<i>Прогнозирование временных рядов в R. Выбор и комплексное применение методов и инструментальных средств систем искусственного интеллекта, критериев их выбора решение задач по руководству коллективной проектной деятельностью для создания, поддержки и использования комплексных систем на основе аналитики больших данных со стороны заказчика. Выявление небольших по масштабу проектов аналитики, которые потенциально могут представлять интерес для ряда подразделений / служб или для организации в целом Выявление областей деловой деятельности, которые потенциально могут получить отдачу от аналитики.</i>

3.	Визуализация в R	<p>Визуализация данных с помощью R. <i>Выбор и комплексное применение методов и инструментальных средств систем искусственного интеллекта, критериев их выбора решение задач по руководству коллективной проектной деятельностью для создания, поддержки и использования комплексных систем на основе аналитики больших данных со стороны заказчика.</i></p> <p><i>Выявление небольших по масштабу проектов аналитики, которые потенциально могут представлять интерес для ряда подразделений / служб или для организации в целом</i></p> <p><i>Выявление областей деловой деятельности, которые потенциально могут получить отдачу от аналитики.</i></p>
4.	Статистика в R	<p>Классификация и кластеризация в R. <i>Выбор и комплексное применение методов и инструментальных средств систем искусственного интеллекта, критериев их выбора решение задач по руководству коллективной проектной деятельностью для создания, поддержки и использования комплексных систем на основе аналитики больших данных со стороны заказчика.</i></p> <p><i>Выявление небольших по масштабу проектов аналитики, которые потенциально могут представлять интерес для ряда подразделений / служб или для организации в целом</i></p> <p><i>Выявление областей деловой деятельности, которые потенциально могут получить отдачу от аналитики.</i></p>

5.2.4. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине

Очная форма обучения

№	Наименование раздела дисциплины	Содержание	Учебно-методическое обеспечение
1	2	3	4
1.	Базовые структуры и понятия языка R	Изучение теоретического материала по рекомендованной в рабочей программе литературе и методического материала, размещенного на образовательном портале АГАСУ. Подготовка к контрольной работе, итоговому тестированию, экзамену.	[3,9,12, 14-16]
2.	Предобработка данных в R		[2,5,8, 14-16]
3.	Визуализация в R		[1,9,10, 14-16]
4.	Статистика в R		[4,6,7,11,13, 14-16]

Заочная форма обучения

№	Наименование раздела дисциплины	Содержание	Учебно-методическое обеспечение
1	2	3	4
1.	Базовые структуры и понятия языка R	Изучение теоретического материала по рекомендованной в рабочей программе литературе и методического материала, размещенного на образовательном портале АГАСУ. Подготовка к контрольной работе, итоговому тестированию, экзамену.	[3,9,12, 14-16]
2.	Предобработка данных в R		[2,5,8, 14-16]
3.	Визуализация в R		[1,9,10, 14-16]
4.	Статистика в R		[4,6,7,11,13, 14-16]

5.2.5. Темы контрольных работ

1. Базовые структуры и понятия языка R
2. Предобработка данных в R
3. Визуализация в R
4. Статистика в R

5.2.6. Темы курсовых проектов/ курсовых работ

Учебным планом не предусмотрены.

6 Методические указания для обучающихся по освоению дисциплины

Организация деятельности студента
<p>Лекция. В ходе лекционных занятий необходимо вести конспектирование учебного материала, обращать внимание на категории, формулировки, раскрывающие содержание тех или иных явлений и процессов, научные выводы и практические рекомендации. Необходимо задавать преподавателю уточняющие вопросы с целью уяснения теоретических положений, разрешения спорных ситуаций. Целесообразно дорабатывать свой конспект лекции, делая в нем соответствующие записи из литературы, рекомендованной преподавателем и предусмотренной учебной программой.</p>
<p>Практическое занятие. Целью практических занятий является углубление и закрепление теоретических знаний, полученных студентами на лекциях и в процессе самостоятельного изучения учебного материала, а, следовательно, формирование у них определенных умений и навыков. На практических занятиях студент вначале знакомится с содержанием работы, пользуясь электронными методическими материалами, размещенными на образовательном портале АГАСУ, затем выполняет задание и показывает результаты преподавателю. Практические работы, выполняются студентом самостоятельно, возникающие при их выполнении проблемы разрешаются в рамках учебного времени и индивидуальных и групповых консультаций.</p>

Самостоятельная работа. Самостоятельная работа студента над усвоением учебного материала по учебной дисциплине может выполняться в помещениях для самостоятельной работы, а также в домашних условиях. Содержание самостоятельной работы студента определяется учебной программой дисциплины, методическими материалами, заданиями и указаниями преподавателя. Самостоятельная работа в аудиторное время включает работу с учебно-методическим материалом, выполнение заданий практических работ.

Самостоятельная работа во внеаудиторное время может состоять из:

- повторения теоретического материала;
- подготовки к практическим занятиям;
- изучения учебной и научной литературы;
- подготовка к итоговому тестированию;
- самопроверка изученного учебного материала.

Контрольная работа.

Теоретическая и практическая части контрольной работы выполняются по установленным темам (вариантам) с использованием практических материалов, полученных на практических (лабораторных) занятиях и при прохождении практики. К каждой теме контрольной работы рекомендуется примерный перечень основных вопросов, список необходимой литературы. Необходимо изучить литературу, рекомендуемую для выполнения контрольной работы. Чтобы полнее раскрыть тему, следует использовать дополнительные источники и материалы. Инструкция по выполнению контрольной работы находится в методических материалах по дисциплине.

Подготовка к экзамену. Подготовка студентов к экзамену включает:

- самостоятельная работа в течение семестра;
- непосредственная подготовка в дни, предшествующие экзамену.

7 Образовательные технологии

Перечень образовательных технологий, используемых при изучении дисциплины «Системы обработки больших данных».

Традиционные образовательные технологии

Дисциплина «Системы обработки больших данных» проводится с использованием традиционных образовательных технологий, ориентирующихся на организацию образовательного процесса, предполагающую прямую трансляцию знаний от преподавателя к студенту (преимущественно на основе объяснительно-иллюстративных методов обучения), учебная деятельность студента носит в таких условиях, как правило, репродуктивный характер. Формы учебных занятий с использованием традиционных технологий:

Лекция – последовательное изложение материала в дисциплинарной логике, осуществляемое преимущественно вербальными средствами (монолог преподавателя).

Практическое занятие – занятие, посвященное освоению конкретных умений и навыков по предложенному алгоритму.

Интерактивные технологии

По дисциплине «Системы обработки больших данных» лекционные занятия проводятся с использованием следующих интерактивных технологий:

Лекция-визуализация представляет собой визуальную форму подачи лекционного материала средствами ТСО или аудиовидеотехники (видео-лекция). Чтение такой лекции сводится к развернутому или краткому комментированию просматриваемых визуальных материалов (в виде схем, таблиц, графов, графиков, моделей). Лекция-визуализация помогает студентам преобразовывать лекционный материал в визуальную

форму, что способствует формированию у них профессионального мышления за счет систематизации и выделения наиболее значимых, существенных элементов.

По дисциплине «Системы обработки больших данных» практические занятия проводятся с использованием следующих интерактивных технологий:

Работа в малых группах – это одна из самых популярных стратегий, так как она дает всем обучающимся (в том числе и стеснительным) возможность участвовать в работе, практиковать навыки сотрудничества, межличностного общения (в частности, умение активно слушать, вырабатывать общее мнение, разрешать возникающие разногласия). Все это часто бывает невозможно в большом коллективе.

Разработка проекта (метод проектов) – организация обучения, при которой учащиеся приобретают знания в процессе планирования и выполнения практических заданий-проектов.

8 Учебно-методическое и информационное обеспечение дисциплины

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

а) основная учебная литература:

1. Черемухин, А. Д. Большие данные: учебное пособие / А. Д. Черемухин. – Москва: Ай Пи Ар Медиа, 2023. – 782 с. – ISBN 978-5-4497-2138-9. – Текст: электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. – URL: <https://www.iprbookshop.ru/129721.html>
2. Целых, А. Н. Применение временных рядов для анализа больших данных: учебное пособие по курсу «Математические методы анализа больших данных» / А. Н. Целых, В. С. Васильев, Э. М. Котов. – Ростов-на-Дону, Таганрог: Издательство Южного федерального университета, 2021. – 84 с. – ISBN 978-5-9275-3983-3. – Текст: электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. – URL: <https://www.iprbookshop.ru/121929.html>
3. Целых А.Н. Современные методы прикладной информатики в задачах анализа данных: учебное пособие по курсу «Методы интеллектуального анализа данных» / Целых А.Н., Целых А.А., Котов Э.М. – Ростов-на-Дону, Таганрог: Издательство Южного федерального университета, 2021. – 130 с. – ISBN 978-5-9275-3783-9. – Текст: электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. – URL: <https://www.iprbookshop.ru/117165.html>

б) дополнительная учебная литература:

4. Железнов М.М. Методы и технологии обработки больших данных: учебно-методическое пособие / Железнов М.М. – Москва: МИСИ-МГСУ, ЭБС АСВ, 2020. – 46 с. – ISBN 978-5-7264-2193-3. – Текст : электронный // Цифровой образовательный ресурс IPR SMART : [сайт]. – URL: <https://www.iprbookshop.ru/101802.html>
5. Воронов В.И. Data Mining - технологии обработки больших данных: учебное пособие / Воронов В.И., Воронова Л.И., Усачев В.А. – Москва: Московский технический университет связи и информатики, 2018. – 47 с. – Текст: электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. – URL: <https://www.iprbookshop.ru/81324.html>
6. Адлер, Ю. П. Статистическое управление процессами. «Большие данные»: учебное пособие / Ю. П. Адлер, Е. А. Черных. – Москва: Издательский Дом МИСиС, 2016. – 52 с. – ISBN 978-5-87623-969-3. – Текст: электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. – URL: <https://www.iprbookshop.ru/64199.html>

в) перечень учебно-методического обеспечения

7. Кравченя П.Д. Системы обработки больших данных: учеб. пособие / П.Д. Кравченя, А.Е. Андреев ВолГГТУ. – Волгоград, 2021. – 40 с.

г) перечень онлайн курсов:

8. ЭОР «Анализ данных в R. Часть 2». Институт биоинформатики.
<https://stepik.org/course/724/syllabus?search=4537986503>

9. ЭОР «Введение в Data Science и машинное обучение». Институт биоинформатики.
<https://stepik.org/course/4852/info>

8.2. Перечень необходимого лицензионного и свободно распространяемого программного обеспечения, в том числе отечественного производства, используемого при осуществлении образовательного процесса по дисциплине

- 7-Zip;
- Adobe Acrobat Reader DC;
- Apache Open Office;
- VLC media player;
- Kaspersky Endpoint Security
- Yandex browser
- R

8.3. Перечень современных профессиональных баз данных и информационных справочных систем, доступных обучающимся при освоении дисциплины

1. Электронная информационно-образовательная среда Университета: образовательный портал (<http://moodle.aucu.ru>)

2. Электронно-библиотечная система «Университетская библиотека» (<https://biblioclub.ru/>)

3. Электронно-библиотечная система «IPRbooks» (www.iprbookshop.ru)

4. Научная электронная библиотека (<http://www.elibrary.ru/>)

9 Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине

№ п/п	Наименование специальных помещений и помещений для самостоятельной работы	Оснащенность специальных помещений и помещений для самостоятельной работы
1.	Учебные аудитории для проведения учебных занятий: 414056, г. Астрахань, ул. Татищева, 18, аудитории № 204, 207, 209, 211	№204 Комплект учебной мебели Учебно-наглядные пособия Стационарный мультимедийный комплект Доступ к информационно – телекоммуникационной сети «Интернет»
		№207 Комплект учебной мебели Компьютеры - 15 шт. Стационарный мультимедийный комплект

№ п/п	Наименование специальных помещений и помещений для самостоятельной работы	Оснащенность специальных помещений и помещений для самостоятельной работы
		<p>Доступ к информационно – телекоммуникационной сети «Интернет»</p> <p>№209 Комплект учебной мебели Компьютеры - 15 шт. Стационарный мультимедийный комплект Графические планшеты – 16 шт. Источник бесперебойного питания – 1шт. Доступ к информационно – телекоммуникационной сети «Интернет»</p> <p>№211 Комплект учебной мебели Компьютеры - 15 шт. Стационарный мультимедийный комплект Доступ к информационно – телекоммуникационной сети «Интернет»</p>
2.	<p>Помещения для самостоятельной работы:</p> <p>414056, г. Астрахань, ул. Татищева, 22а, аудитории №201,203</p> <p>414056, г. Астрахань, ул. Татищева, 18а библиотека, читальный зал</p>	<p>№201 Комплект учебной мебели Компьютеры – 8 шт. Доступ к информационно – телекоммуникационной сети «Интернет»</p> <p>№203 Комплект учебной мебели Компьютеры – 8 шт. Доступ к информационно – телекоммуникационной сети «Интернет»</p> <p>библиотека, читальный зал Комплект учебной мебели Компьютеры - 4 шт. Доступ к информационно – телекоммуникационной сети «Интернет»</p>

10. Особенности организации обучения по дисциплине для инвалидов и лиц с ограниченными возможностями здоровья

Для обучающихся из числа инвалидов и лиц с ограниченными возможностями здоровья на основании письменного заявления дисциплина «Системы обработки больших данных» реализуется с учетом особенностей психофизического развития, индивидуальных возможностей и состояния здоровья (далее – индивидуальных особенностей).

Аннотация

к рабочей программе дисциплины «Системы обработки больших данных»
по направлению подготовки 09.04.02 «Информационные системы и технологии»
направленность (профиль) «Искусственный интеллект в проектировании городской
среды»

*Общая трудоемкость дисциплины составляет 3 зачетные единицы.
Форма промежуточной аттестации: экзамен.*

Целью освоения учебной дисциплины «Системы обработки больших данных» является формирование компетенций обучающихся в соответствии с требованиями Федерального государственного образовательного стандарта высшего образования по направлению подготовки 09.04.02 «Информационные системы и технологии».

Учебная дисциплина Б1.О.6 «Системы обработки больших данных» входит в Блок 1 «Дисциплины (модули)» обязательной части. Для освоения дисциплины необходимы знания, полученные при изучении следующих дисциплин: «Специальные главы математики», «Технологии программирования и инструментальные средства разработки систем искусственного интеллекта», «Технологии анализа данных».

Краткое содержание дисциплины

- Раздел 1. Базовые структуры и понятия языка R
- Раздел 2. Предобработка данных в R
- Раздел 3. Визуализация в R
- Раздел 4. Статистика в R

И.о. заведующего кафедрой



(подпись)

/В.В. Соболева/
И.О.Ф

РЕЦЕНЗИЯ
на рабочую программу, оценочные и методические материалы
Б1.О.6 «Системы обработки больших данных»
(наименование дисциплины с указанием блока)

ОПОП ВО по направлению подготовки
09.04.02 «Информационные системы и технологии»
направленность (профиль)
«Искусственный интеллект в проектировании городской среды»
по программе магистратуры

Соболева В.В. (далее по тексту рецензент), проведена рецензия рабочей программы, оценочных и методических материалов по дисциплине «Системы обработки больших данных» ОПОП ВО по направлению подготовки 09.04.02 «Информационные системы и технологии», направленность (профиль) «Искусственный интеллект в проектировании городской среды», по программе *магистратуры*, разработанной в ГБОУ АО ВО "Астраханский государственный архитектурно-строительный университет", на кафедре САПРиМ (разработчик – Окладникова С.В.).

Рассмотрев представленные на рецензию материалы, рецензент пришел к следующим выводам:

Предъявленная рабочая программа учебной дисциплины «Системы обработки больших данных» (далее по тексту Программа) соответствует требованиям ФГОС ВО по направлению подготовки 09.04.02 «Информационные системы и технологии», утвержденного приказом Министерства образования и науки Российской Федерации от 19.09.2017г. №917, редакция с изменениями № 1456 от 26.11.2020г., с изменениями и дополнениями от 26.11.2020 г., 08.02.2021 г. и зарегистрированного в Минюсте России от 16.10.2017г, №48550.

Представленная в Программе актуальность учебной дисциплины в рамках реализации ОПОП ВО не подлежит сомнению – дисциплина относится к Блоку 1. «Дисциплины (модули)» обязательной части.

Представленные в Программе цели учебной дисциплины соответствуют требованиям ФГОС ВО по направлению подготовки 09.04.02 «Информационные системы и технологии», направленность (профиль) «Искусственный интеллект в проектировании городской среды».

В соответствии с Программой за дисциплиной «Системы обработки больших данных» закреплены 2 компетенции, которые реализуются в объявленных требованиях.

Предложенные в Программе индикаторы компетенций в категориях знать, уметь, иметь практический опыт отражают специфику и содержание дисциплины, а представленные в ОММ показатели и критерии оценивания компетенций по дисциплине на различных этапах их формирования, а также шкалы оценивания позволяют определить степень достижения заявленных результатов, т.е. уровень освоения обучающимися соответствующих компетенций в рамках дисциплины.

Учебная дисциплина «Системы обработки больших данных» взаимосвязана с другими дисциплинами ОПОП ВО по направлению подготовки 09.04.02 «Информационные системы и технологии», направленность (профиль) «Искусственный интеллект в проектировании городской среды» и возможность дублирования в содержании не выявлена.

Представленная Программа предполагает использование современных образовательных технологий при реализации различных видов учебной работы. Формы образовательных технологий соответствуют специфике дисциплины.

Представленные и описанные в Программе формы текущей оценки знаний соответствуют специфике дисциплины и требованиям к выпускникам.

Промежуточная аттестация знаний *магистратуры*, предусмотренная Программой, осуществляется в форме *экзамена*. Формы оценки знаний, представленные в Рабочей программе, соответствуют специфике дисциплины и требованиям к выпускникам.

Учебно-методическое обеспечение дисциплины представлено основной, дополнительной литературой, интернет-ресурсами и соответствует требованиям ФГОС ВО

РЕЦЕНЗИЯ
на рабочую программу, оценочные и методические материалы
Б1.О.6 «Системы обработки больших данных»
(наименование дисциплины с указанием блока)

ОПОП ВО по направлению подготовки
09.04.02 «Информационные системы и технологии»
направленность (профиль)
«Искусственный интеллект в проектировании городской среды»
по программе магистратуры

Хоменко Т.В. (далее по тексту рецензент), проведена рецензия рабочей программы, оценочных и методических материалов по дисциплине «Системы обработки больших данных» ОПОП ВО по направлению подготовки 09.04.02 «Информационные системы и технологии», направленность (профиль) «Искусственный интеллект в проектировании городской среды», по программе *магистратуры*, разработанной в ГБОУ АО ВО "Астраханский государственный архитектурно-строительный университет", на кафедре САПРиМ (разработчик – Окладникова С.В.).

Рассмотрев представленные на рецензию материалы, рецензент пришел к следующим выводам:

Предъявленная рабочая программа учебной дисциплины «Системы обработки больших данных» (далее по тексту Программа) соответствует требованиям ФГОС ВО по направлению подготовки 09.04.02 «Информационные системы и технологии», утвержденного приказом Министерства образования и науки Российской Федерации от 19.09.2017г. №917, редакция с изменениями № 1456 от 26.11.2020г., с изменениями и дополнениями от 26.11.2020 г., 08.02.2021 г. и зарегистрированного в Минюсте России от 16.10.2017г, №48550.

Представленная в Программе актуальность учебной дисциплины в рамках реализации ОПОП ВО не подлежит сомнению – дисциплина относится к Блоку 1 «Дисциплины (модули)» обязательной части.

Представленные в Программе цели учебной дисциплины соответствуют требованиям ФГОС ВО по направлению подготовки 09.04.02 «Информационные системы и технологии», направленность (профиль) «Искусственный интеллект в проектировании городской среды».

В соответствии с Программой за дисциплиной «Системы обработки больших данных» закреплены 2 компетенции, которые реализуются в объявленных требованиях.

Предложенные в Программе индикаторы компетенций в категориях знать, уметь, иметь практический опыт отражают специфику и содержание дисциплины, а представленные в ОММ показатели и критерии оценивания компетенций по дисциплине на различных этапах их формирования, а также шкалы оценивания позволяют определить степень достижения заявленных результатов, т.е. уровень освоения обучающимися соответствующих компетенций в рамках дисциплины.

Учебная дисциплина «Системы обработки больших данных» взаимосвязана с другими дисциплинами ОПОП ВО по направлению подготовки 09.04.02 «Информационные системы и технологии», направленность (профиль) «Искусственный интеллект в проектировании городской среды» и возможность дублирования в содержании не выявлена.

Представленная Программа предполагает использование современных образовательных технологий при реализации различных видов учебной работы. Формы образовательных технологий соответствуют специфике дисциплины.

Представленные и описанные в Программе формы текущей оценки знаний соответствуют специфике дисциплины и требованиям к выпускникам.

Промежуточная аттестация знаний *магистратуры*, предусмотренная Программой, осуществляется в форме *экзамена*. Формы оценки знаний, представленные в Рабочей программе, соответствуют специфике дисциплины и требованиям к выпускникам.

Учебно-методическое обеспечение дисциплины представлено основной, дополнительной литературой, интернет-ресурсами и соответствует требованиям ФГОС ВО

направления подготовки 09.04.02 «Информационные системы и технологии», направленность (профиль) «Искусственный интеллект в проектировании городской среды».

Материально-техническое обеспечение соответствует требованиям ФГОС ВО направления подготовки 09.04.02 «Информационные системы и технологии», направленность (профиль) «Искусственный интеллект в проектировании городской среды» и специфике дисциплины «Системы обработки больших данных» и обеспечивает использование современных образовательных, в том числе интерактивных методов обучения.

Представленные на рецензию оценочные и методические материалы по направлению подготовки 09.04.02 «Информационные системы и технологии», разработаны в соответствии с нормативными документами, представленными в Программе. Оценочные и методические материалы по дисциплине «Системы обработки больших данных» предназначены для текущего контроля и промежуточной аттестации и представляет собой совокупность разработанных кафедрой САПРиМ материалов для установления уровня и качества достижения обучающимися результатов обучения.

Задачами оценочных и методических материалов является контроль и управление процессом освоения обучающимися компетенций, заявленных в образовательной программе по направлению подготовки 09.04.02 «Информационные системы и технологии», направленность (профиль) «Искусственный интеллект в проектировании городской среды».

Оценочные и методические материалы по дисциплине «Системы обработки больших данных» представлены: перечнем материалов текущего контроля и промежуточной аттестации.

Данные материалы позволяют в полной мере оценить результаты обучения по дисциплине «Системы обработки больших данных» в АГАСУ, а также оценить степень сформированной компетенций.

ОБЩИЕ ВЫВОДЫ

На основании проведенной рецензии можно сделать заключение, что характер, структура, содержание рабочей программы, оценочных и методических материалов дисциплины «Системы обработки больших данных» ОПОП ВО по направлению подготовки 09.04.02 «Информационные системы и технологии», направленность (профиль) «Искусственный интеллект в проектировании городской среды», по программе *магистратуры*, разработанная к.т.н., доц. Окладниковой С.В. соответствует требованиям ФГОС ВО, современным требованиям отрасли, рынка труда, профессиональных стандартов направлению подготовки 09.04.02 «Информационные системы и технологии», направленность (профиль) «Искусственный интеллект в проектировании городской среды» и могут быть рекомендованы к использованию.

Рецензент:

Хоменко Татьяна Владимировна,
доктор технических наук, доцент
зав. кафедрой «Автоматизированные
системы обработки информации и
управления (АСОИУ)» ФГБОУ ВО
«Астраханский государственный
технический университет»


(подпись)

/ Т.В. Хоменко/
(И.О.Ф.)



Министерство образования и науки Астраханской области
Государственное бюджетное образовательное учреждение
Астраханской области высшего образования
«Астраханский государственный архитектурно-строительный
университет»
(ГБОУ АО ВО «АГАСУ»)

УТВЕРЖДАЮ
И.о. первого проректора
С.П. Стрелков /
И. О. Ф.
2024 г.



ОЦЕНОЧНЫЕ И МЕТОДИЧЕСКИЕ МАТЕРИАЛЫ

Наименование дисциплины

«Системы обработки больших данных»

(указывается наименование в соответствии с учебным планом)

По направлению подготовки

09.04.02 «Информационные системы и технологии»

(указывается наименование направления подготовки в соответствии с ФГОС ВО)

Направленность (профиль)

«Искусственный интеллект в проектировании городской среды»

(указывается наименование направленности (профиля) в соответствии с ОПОП)

Кафедра

Системы автоматизированного проектирования и моделирования

Квалификация выпускника *магистр*

Разработчик:

К.т.н., доцент

(занимаемая должность,
учёная степень и учёное звание)



(подпись)

/С.В. Окладникова/

(И. О. Ф.)

Оценочные и методические материалы рассмотрены и утверждены на заседании кафедры
«Системы автоматизированного проектирования и моделирования»

Протокол № 9 от 22.04 2022 г.

И.о. заведующего кафедрой



(подпись)

/В.В.Соболева/

(И. О. Ф.)

Согласовано:

Председатель МКН «Информационные системы и технологии»

направленность(профиль) «Искусственный интеллект в проектировании городской среды»



(подпись)

/В.В. Соболева/
И.О.Ф.

Начальник УМУ

 | А.С. Степанов
(подпись) (И. О. Ф)

Специалист УМУ

 | А.С. Степанов |
(подпись) (И. О. Ф)

Содержание

1. Оценочные и методические материалы для проведения текущего контроля успеваемости и промежуточной аттестации обучающихся по дисциплине	4
1.1. Перечень компетенций с указанием этапов их формирования в процессе освоения образовательной программы.....	4
1.2. Описание показателей и критериев оценивания компетенций на различных этапах их формирования, описание шкал оценивания.....	5
1.2.1. Перечень оценочных средств текущего контроля успеваемости.....	5
1.2.2. Описание показателей и критериев оценивания компетенций по дисциплине на различных этапах их формирования, описание шкал оценивания	5
1.2.3. Шкала оценивания	9
2. Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков, характеризующих этапы формирования компетенций в процессе освоения образовательной программы.....	10
2.1. Экзамен	10
2.2. Защита практической работы	11
2.3. Контрольная работа	11
2.4. Тест.....	12
3. Перечень и характеристики процедуры оценивания знаний, умений, навыков, характеризующих этапы формирования компетенций	13
<i>Приложение 1</i>	14
<i>Приложение 2</i>	15
<i>Приложение 3</i>	25
<i>Приложение 4</i>	28
<i>Приложение 5</i>	31

1. Оценочные и методические материалы для проведения текущего контроля успеваемости и промежуточной аттестации обучающихся по дисциплине

Оценочные и методические материалы являются неотъемлемой частью рабочей программы дисциплины (далее РПД) и представлены в виде отдельного документа.

1.1. Перечень компетенций с указанием этапов их формирования в процессе освоения образовательной программы

Индекс и формулировка компетенции		Индикаторы достижений компетенций, установленные ОПОП	Номер раздела дисциплины (в соответствии с п.5.1 РПД)				Формы контроля с конкретизацией задания
			1	2	3	4	
1		2	3	4	5	6	7
ПК-1ИИП. Способен исследовать применение интеллектуальных систем для различных предметных областей	ПК-1ИИП.2 – Выбирает комплексы методов и инструментальных средств искусственного интеллекта для решения задач в зависимости от особенностей предметной области	Знать: методы и инструментальные средства систем искусственного интеллекта, критерии их выбора и методы комплексирования в рамках применения интегрированных гибридных интеллектуальных систем различного назначения (ПК-1ИИП.2 З-1)	X	X	X	X	Вопросы к экзамену [1-8] Итоговое тестирование [1-23]
		Уметь: выбирать и комплексно применять методы и инструментальные средства систем искусственного интеллекта, критерии их выбора (ПК-1ИИП.2 У-1)	X	X			Контрольная работа [задание 1] Отчет по выполнению практической работы
ПК-4ИИП. Способен руководить проектами по созданию комплексных систем на основе аналитики	ПК4.1ИИП – Руководит проектами по построению комплексных систем на основе аналитики больших	Знать: Знает методологию и принципы руководства проектами по созданию, поддержке и использованию комплексных систем на основе аналитики больших данных со стороны заказчика (ПК-4.1ИИП. З-1.) Знает специфику сфер и отраслей, для которых реализуется проект по аналитике больших данных (ПК-4.1ИИП. З-2.)	X	X	X	X	Вопросы к экзамену [9-17] Итоговое тестирование [24-40] Вопросы к экзамену [18-26] Итоговое тестирование [41-63]

больших данных в различных отраслях со стороны заказчика	данных в различных отраслях со стороны заказчика	Уметь:					
		<p>Умеет решать задачи по руководству коллективной проектной деятельностью для создания, поддержки и использования комплексных систем на основе аналитики больших данных со стороны заказчика (ПК-4.1ИИП. У-1.)</p> <p>Умеет выявлять небольшие по масштабу проекты аналитики, которые потенциально могут представлять интерес для ряда подразделений / служб или для организации в целом (ПК-4.1ИИП. У-2.)</p> <p>Умеет выявлять области деловой деятельности, которые потенциально могут получить отдачу от аналитики (ПК-4.1ИИП. У-3.)</p>			X	X	Контрольная работа [задания 2, 3] Отчет по выполнению практической работы

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

1.2.1. Перечень оценочных средств текущего контроля успеваемости

Наименование оценочного средства	Краткая характеристика оценочного средства	Представление оценочного средства в фонде
Практическая работа	Средство, позволяющее оценить умение и владение обучающегося излагать суть поставленной задачи, самостоятельно применять стандартные методы решения поставленной задачи с использованием имеющейся лабораторно-практической базы, проводить анализ полученного результата работы. Рекомендуется для оценки умений и владений студентов	Темы практических работ и требования к их защите
Контрольная работа	Средство проверки умений применять полученные знания для решения задач определенного типа по теме или разделу	Задания контрольной работы
Тест	Система стандартизированных заданий, позволяющая автоматизировать процедуру измерения уровня знаний обучающегося	Фонд тестовых заданий

1.2.2. Описание показателей и критериев оценивания компетенций по дисциплине на различных этапах их формирования, описание шкал оценивания

Компетенция, этапы освоения компетенции		Планируемые результаты обучения	Показатели и критерии оценивания результатов обучения			
			Ниже порогового уровня (неудовлетворитель но)	Пороговый уровень (удовлетворитель но)	Продвинутый уровень (хорошо)	Высокий уровень (отлично)
1		2	3	4	5	6
ПК-1ИИП. Способен исследовать применение интеллектуаль ных систем для различных предметных областей	ПК-1ИИП.2 – Выбирает комплексы методов и инструменталь ных средств искусственного интеллекта для решения задач в зависимости от особенностей предметной области	Знать: ПК-1ИИП.2 З-1. Знает методы и инструментальны е средства систем искусственного интеллекта, критерии их выбора и методы комплексировани я в рамках применения интегрированных гибридных интеллектуальных систем различного назначения	Не знает методы и инструментальные средства систем искусственного интеллекта, критерии их выбора и методы комплексирования в рамках применения интегрированных гибридных интеллектуальных систем различного назначения	Знает методы и инструментальные средства систем искусственного интеллекта, критерии их выбора и методы комплексирования в рамках применения интегрированных гибридных интеллектуальных систем различного назначения	Знает методы и инструментальн ые средства систем искусственного интеллекта, критерии их выбора и методы комплексирован ия в рамках применения интегрированны х гибридных интеллектуальн ых систем различного назначения	Знает методы и инструментальн ые средства систем искусственного интеллекта, критерии их выбора и методы комплексирован ия в рамках применения интегрированны х гибридных интеллектуальн ых систем различного назначения
		Уметь: ПК-1ИИП.2 У-1. Умеет выбирать и комплексно применять методы и инструментальны е средства систем искусственного интеллекта, критерии их выбора	Не умеет выбирать и комплексно применять методы и инструментальные средства систем искусственного интеллекта, критерии их выбора	Умеет выбирать и комплексно применять методы и инструментальные средства систем искусственного интеллекта, критерии их выбора	Умеет выбирать и комплексно применять методы и инструментальн ые средства систем искусственного интеллекта, критерии их выбора	Умеет выбирать и комплексно применять методы и инструментальн ые средства систем искусственного интеллекта, критерии их выбора

Компетенция, этапы освоения компетенции		Планируемые результаты обучения	Показатели и критерии оценивания результатов обучения			
			Ниже порогового уровня (неудовлетворитель но)	Пороговый уровень (удовлетворитель но)	Продвинутый уровень (хорошо)	Высокий уровень (отлично)
1		2	3	4	5	6
ПК-4ИИП. Способен руководить проектами по созданию комплексных систем на основе аналитики больших данных в различных отраслях со стороны заказчика	ПК4.1ИИП – Руководит проектами по построению комплексных систем на основе аналитики больших данных в различных отраслях со стороны заказчика	Знать: Знает методологию и принципы руководства проектами по созданию, поддержке и использованию комплексных систем на основе аналитики больших данных со стороны заказчика (ПК- 4.1ИИП. 3-1.) Знает специфику сфер и отраслей, для которых реализуется проект по аналитике больших данных (ПК-4.1ИИП. 3- 2.)	Не знает методологию и принципы руководства проектами по созданию, поддержке и использованию комплексных систем на основе аналитики больших данных со стороны заказчика. Не знает специфику сфер и отраслей, для которых реализуется проект по аналитике больших данных.	Знает методологию и принципы руководства проектами по созданию, поддержке и использованию комплексных систем на основе аналитики больших данных со стороны заказчика. Знает специфику сфер и отраслей, для которых реализуется проект по аналитике больших данных.	Знает методологию и принципы руководства проектами по созданию, поддержке и использованию комплексных систем на основе аналитики больших данных со стороны заказчика. Знает специфику сфер и отраслей, для которых реализуется проект по аналитике больших данных.	Знает методологию и принципы руководства проектами по созданию, поддержке и использованию комплексных систем на основе аналитики больших данных со стороны заказчика. Знает специфику сфер и отраслей, для которых реализуется проект по аналитике больших данных.

Компетенция, этапы освоения компетенции	Планируемые результаты обучения	Показатели и критерии оценивания результатов обучения			
		Ниже порогового уровня (неудовлетворитель но)	Пороговый уровень (удовлетворитель но)	Продвинутый уровень (хорошо)	Высокий уровень (отлично)
1	2	3	4	5	6
	<p>Уметь: Умеет решать задачи по руководству коллективной проектной деятельностью для создания, поддержки и использования комплексных систем на основе аналитики больших данных со стороны заказчика (ПК-4.1ИИП. У-1.) Умеет выявлять небольшие по масштабу проекты аналитики, которые потенциально могут представлять интерес для ряда подразделений / служб или для организации в целом (ПК-</p>	<p>Не умеет решать задачи по руководству коллективной проектной деятельностью для создания, поддержки и использования комплексных систем на основе аналитики больших данных со стороны заказчика. Не умеет выявлять небольшие по масштабу проекты аналитики, которые потенциально могут представлять интерес для ряда подразделений / служб или для организации в целом. Не умеет выявлять области деловой деятельности, которые потенциально могут получить отдачу от аналитики.</p>	<p>Умеет решать задачи по руководству коллективной проектной деятельностью для создания, поддержки и использования комплексных систем на основе аналитики больших данных со стороны заказчика. Умеет выявлять небольшие по масштабу проекты аналитики, которые потенциально могут представлять интерес для ряда подразделений / служб или для организации в целом. Умеет выявлять области деловой деятельности, которые потенциально могут получить</p>	<p>Умеет решать задачи по руководству коллективной проектной деятельностью для создания, поддержки и использования комплексных систем на основе аналитики больших данных со стороны заказчика. Умеет выявлять небольшие по масштабу проекты аналитики, которые потенциально могут представлять интерес для ряда подразделений / служб или для организации в целом.</p>	<p>Умеет решать задачи по руководству коллективной проектной деятельностью для создания, поддержки и использования комплексных систем на основе аналитики больших данных со стороны заказчика. Умеет выявлять небольшие по масштабу проекты аналитики, которые потенциально могут представлять интерес для ряда подразделений / служб или для организации в целом.</p>

Компетенция, этапы освоения компетенции	Планируемые результаты обучения	Показатели и критерии оценивания результатов обучения			
		Ниже порогового уровня (неудовлетворитель но)	Пороговый уровень (удовлетворитель но)	Продвинутый уровень (хорошо)	Высокий уровень (отлично)
1	2	3	4	5	6
	4.1ИИП. У-2.) Умеет выявлять области деловой деятельности, которые потенциально могут получить отдачу от аналитики (ПК- 4.1ИИП. У-3.)		отдачу от аналитики.	Умеет выявлять области деловой деятельности, которые потенциально могут получить отдачу от аналитики.	Умеет выявлять области деловой деятельности, которые потенциально могут получить отдачу от аналитики.

1.2.3. Шкала оценивания

Уровень достижений	Отметка в 5-бальной шкале
высокий	«5»(отлично)
продвинутый	«4»(хорошо)
пороговый	«3»(удовлетворительно)
ниже порогового	«2»(неудовлетворительно)

2. Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков, характеризующих этапы формирования компетенций в процессе освоения образовательной программы

ТИПОВЫЕ ЗАДАНИЯ ДЛЯ ПРОВЕДЕНИЯ ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ

2.1. Экзамен

а) типовые вопросы к экзамену (см. приложение 1);

б) критерии оценки.

При оценке знаний на зачете учитывается:

1. Уровень сформированности компетенций.
2. Уровень усвоения теоретических положений дисциплины, правильность формулировки основных понятий и закономерностей.
3. Уровень знания фактического материала в объеме программы.
4. Логика, структура и грамотность изложения вопроса.
5. Умение связать теорию с практикой.
6. Умение делать обобщения, выводы.

№ п/п	Оценка	Критерии оценки
1	отлично	Ответы на поставленные вопросы излагаются логично, последовательно и не требуют дополнительных пояснений. Полно раскрываются причинно-следственные связи между явлениями и событиями. Делаются обоснованные выводы. Демонстрируются глубокие знания базовых нормативно-правовых актов. Соблюдаются нормы литературной речи.
2	хорошо	Ответы на поставленные вопросы излагаются систематизировано и последовательно. Базовые нормативно-правовые акты используются, но в недостаточном объеме. Материал излагается уверенно. Раскрыты причинно-следственные связи между явлениями и событиями. Демонстрируется умение анализировать материал, однако не все выводы носят аргументированный и доказательный характер. Соблюдаются нормы литературной речи.
3	удовлетворительно	Допускаются нарушения в последовательности изложения. Имеются упоминания об отдельных базовых нормативно-правовых актах. Неполно раскрываются причинно-следственные связи между явлениями и событиями. Демонстрируются поверхностные знания вопроса, с трудом решаются конкретные задачи. Имеются затруднения с выводами. Допускаются нарушения норм литературной речи.
4	неудовлетворительно	Материал излагается непоследовательно, сбивчиво, не представляет определенной системы знаний по дисциплине. Не раскрываются причинно-следственные связи между явлениями и событиями. Не проводится анализ. Выводы отсутствуют. Ответы на дополнительные вопросы отсутствуют. Имеются заметные нарушения норм литературной речи.
5	зачтено	Выставляется при соответствии параметрам экзаменационной шкалы на уровнях «отлично», «хорошо», «удовлетворительно».

№ п/п	Оценка	Критерии оценки
6	не зачтено	Выставляется при соответствии параметрам экзаменационной шкалы на уровне «неудовлетворительно».

ТИПОВЫЕ ЗАДАНИЯ ДЛЯ ПРОВЕДЕНИЯ ТЕКУЩЕГО КОНТРОЛЯ

2.2. Защита практической работы

- а) типовые задания практических работ (см. приложение 2);
 б) критерии оценки.

При оценке знаний на защите практической работы учитывается:

1. Уровень сформированности компетенций.
2. Уровень усвоения теоретических положений дисциплины, правильность формулировки основных понятий и закономерностей.
3. Уровень знания фактического материала в объеме программы.
4. Логика, структура и грамотность изложения вопроса.
5. Умение связать теорию с практикой.
6. Умение делать обобщения, выводы.

№ п/п	Оценка	Критерии оценки
1	2	3
1	отлично	Студент правильно называет метод исследования, правильно называет прибор, правильно демонстрирует методику исследования /измерения, правильно оценивает результат.
2	хорошо	Студент правильно называет метод исследования, правильно называет прибор, допускает единичные ошибки в демонстрации методики исследования /измерения и оценке его результатов
3	удовлетворительно	Студент неправильно называет метод исследования, но при этом дает правильное название прибора. Допускает множественные ошибки в демонстрации методики исследования /измерения и оценке его результатов
4	неудовлетворительно	Студент неправильно называет метод исследования, дает неправильное название прибора. Не может продемонстрировать методику исследования /измерения, а также оценить результат
5	зачтено	Выставляется при соответствии параметрам экзаменационной шкалы на уровнях «отлично», «хорошо», «удовлетворительно».
6	не зачтено	Выставляется при соответствии параметрам экзаменационной шкалы на уровне «неудовлетворительно».

2.3. Контрольная работа

- а) варианты контрольной работы (см. приложение 3);
 б) критерии оценки.

При оценке знаний контрольной работы учитывается:

1. Уровень сформированности компетенций.
2. Уровень усвоения теоретических положений дисциплины, правильность формулировки основных понятий и закономерностей.
3. Уровень знания фактического материала в объеме программы.
4. Логика, структура и грамотность изложения вопроса.
5. Умение связать теорию с практикой.

6. Умение делать обобщения, выводы.

№ п/п	Оценка	Критерии оценки
1	2	3
1	отлично	Студент правильно называет метод исследования, правильно называет прибор, правильно демонстрирует методику исследования /измерения, правильно оценивает результат.
2	хорошо	Студент правильно называет метод исследования, правильно называет прибор, допускает единичные ошибки в демонстрации методики исследования /измерения и оценке его результатов
3	удовлетворительно	Студент неправильно называет метод исследования, но при этом дает правильное название прибора. Допускает множественные ошибки в демонстрации методики исследования /измерения и оценке его результатов
4	неудовлетворительно	Студент неправильно называет метод исследования, дает неправильное название прибора. Не может продемонстрировать методику исследования /измерения, а также оценить результат
5	зачтено	Выставляется при соответствии параметрам экзаменационной шкалы на уровнях «отлично», «хорошо», «удовлетворительно».
6	не зачтено	Выставляется при соответствии параметрам экзаменационной шкалы на уровне «неудовлетворительно».

2.4. Тест

- а) *типовой комплект заданий для входного тестирования (Приложение 4)*
типовой комплект заданий для итогового тестирования (Приложение 5)
- б) *критерии оценки.*

При оценке знаний по результатам тестов учитывается:

1. Уровень сформированности компетенций.
2. Уровень усвоения теоретических положений дисциплины, правильность формулировки основных понятий и закономерностей.
3. Уровень знания фактического материала в объеме программы.
4. Логика, структура и грамотность изложения вопроса.
5. Умение связать теорию с практикой.
6. Умение делать обобщения, выводы.

№ п/п	Оценка	Критерии оценки
1	2	3
1	отлично	если выполнены следующие условия: - даны правильные ответы не менее чем на 90% вопросов теста, исключая вопросы, на которые студент должен дать свободный ответ; - на все вопросы, предполагающие свободный ответ, студент дал правильный и полный ответ.
2	хорошо	если выполнены следующие условия: - даны правильные ответы не менее чем на 75% вопросов теста, исключая вопросы, на которые студент должен дать свободный ответ; - на все вопросы, предполагающие свободный ответ, студент дал правильный ответ, но допустил незначительные ошибки и не показал необходимой полноты.

№ п/п	Оценка	Критерии оценки
1	2	3
3	удовлетворительно	если выполнены следующие условия: - даны правильные ответы не менее чем на 50% вопросов теста, исключая вопросы, на которые студент должен дать свободный ответ; -- на все вопросы, предполагающие свободный ответ, студент дал непротиворечивый ответ, или при ответе допустил значительные неточности и не показал полноты.
4	неудовлетворительно	если студентом не выполнены условия, предполагающие оценку «удовлетворительно».
5	зачтено	Выставляется при соответствии параметрам экзаменационной шкалы на уровнях «отлично», «хорошо», «удовлетворительно».
6	не зачтено	Выставляется при соответствии параметрам экзаменационной шкалы на уровне «неудовлетворительно».

3. Перечень и характеристики процедуры оценивания знаний, умений, навыков, характеризующих этапы формирования компетенций

Процедура проведения текущего контроля успеваемости и промежуточной аттестации обучающихся по дисциплине регламентируется локальным нормативным актом.

Перечень и характеристика процедур текущего контроля и промежуточной аттестации по дисциплине

№	Наименование оценочного средства	Периодичность и способ проведения процедуры оценивания	Виды вставляемых оценок	Форма учета
1.	Экзамен	Раз в семестр, по окончании изучения дисциплины	По пятибалльной шкале	Ведомость, зачетная книжка, портфолио
2.	Практическая работа	Систематически на занятиях	По пятибалльной шкале	Лабораторная тетрадь, журнал успеваемости преподавателя
3	Контрольная работа	Один раз в течении семестра	По пятибалльной шкале	журнал успеваемости преподавателя
4	Тест	Входное тестирование по дисциплине – вначале изучения дисциплины (в начале семестра) Итоговое тестирование – по окончании изучения дисциплины	По пятибалльной шкале	Лист результатов компьютерного тестирования, журнал успеваемости преподавателя, ведомость, зачетная книжка, портфолио

**Типовые вопросы к экзамену
по дисциплине «Системы обработки больших данных»**

Знать – ПК-1ИИП.2 3-1

1. Дайте понятие языка R.
2. Раскройте особенности языка R.
3. Назовите отличия языка R от других языков программирования.
4. Расскажите об истории языка программирования R.
5. Раскройте основные особенности установки R и RStudio.
6. Перечислите арифметические операции в R.
7. Перечислите основные типы переменных в R.
8. Особенности массивов в R.

Знать - ПК-4.1ИИП. 3-1

9. Перечислите некоторые функции, которые предоставляет R.
10. Охарактеризуйте формат матриц в R
11. Особенности синтаксиса команд R.
12. Сохранение данных в R.
13. Основные источники и этапы импорта данных в R.
14. Как в R представлены пропущенные значения.
15. Объясните, как данные агрегируются в R.
16. Объясните, как создать таблицу в R без внешнего файла?
17. Принципы работы с API.

Знать - ПК-4.1ИИП. 3-2

18. Доверительные интервалы и t-тест. K-means.
19. Объясните, как можно запустить графический интерфейс R.
20. Какие структуры данных в R используются для статистического анализа?
21. Какие структуры данных в R используются для создания графиков?
22. Объясните, что такое транспонирование.
23. Объясните, что такое t-тесты в R.
24. Для чего используются функции With() и By() в R?
25. Какая функция используется для добавления наборов данных в R?
26. Для чего используются функции subset() и sample() в R?

**Типовые задания практических работ
по дисциплине «Системы обработки больших данных»**

Уметь – ПК-1ИИП.2 У-1, ПК-4.1ИИП. У-1, ПК-4.1ИИП. У-2, ПК-4.1ИИП. У-3

ЛАБОРАТОРНАЯ РАБОТА 1: «Загрузка данных в R»

1. Цель работы

- Приобрести основные навыки работы с данными в R
- Освоить наиболее распространенные способы загрузки данных в рабочую среду R

2. Задачи

- Углубить и закрепить знания по основным типам данных и научить загрузить и изучить свойства данных.
- Совершенствовать навыки самостоятельной работы.
- ...

3. План выполнения работы

3.1 Загрузка различных данных

- Классический вариант: загрузка данных форматов .txt или .csv в среду R: для этого воспользуемся командами:

```
data <- read.csv("путь к файлу/имя файла.csv", sep = ";")
data <- read.table("путь к файлу/имя файла.txt", sep = "\t")
```

- Загрузка файлов в форматах .xls или .xlsx (форматы программы Excel) :

Способ 1: Сохранение excel files в формате .csv, потом использование базовой функции *read.csv*

Способ 2: Используя пакеты **xlsx**, **readxl** **gdata**, **xlsReadWrite**, **XLConnect**,...

```
library(readxl)
M3C <- read_excel("GitHub/LAB/LAB 1/data/M3C.xls")
View(M3C)
```

- Загрузка данных из других пакетов R: к данным доступа из определенного пакета используйте параметр пакета, например:

```
> data(package="Mcomp")
> data(M3, package="Mcomp")
```

- Загрузка данных из интернета: для этого воспользуемся пакетом **Rcurl**

Например: Загрузка набора данных AirPassengers – количество перевозки авиапассажиров 1949-1960.

```
> library(RCurl)
> myfile <-
getURL('https://raw.githubusercontent.com/vincentarelbundock/Rdatasets/master/csv/datasets/AirPassengers.csv',
ssl.verifyhost=FALSE, ssl.verifypeer=FALSE)
> data <- read.csv(textConnection(myfile), sep = ",", header = TRUE)
> head(data)
  X   time AirPassengers
1 1 1949.000          112
2 2 1949.083          118
3 3 1949.167          132
4 4 1949.250          129
5 5 1949.333          121
6 6 1949.417          135
```

3.2 Свойства данных

Предлагаем, что мы уже загрузили данные в виде data frame с именем data1:

```
data1 <- read.csv("FORA.csv", sep = ";")
```

Для получения свойства загруженных данных воспользуемся следующими командами:

- head(x) — посмотреть начальные строки объекта
- tail(x) — посмотреть последние строки объекта
- dput(x) — получить представление объекта в R-синтаксисе
- summary(object) — общая информация об объекте
- str(object) — внутренняя структура объекта
- dim(x) — размерности объекта
- dimnames(x) — имена размерностей объекта
- names(x) — имена объекта
- length(x) — число элементов

Например:

- head(data1)
- summary(data1)

3.3 Работа с Data Frame

Data frame - это гибридный тип представления, одномерный список из векторов одинаковой длины. Таким образом, каждая data frame - это список колонок, причём внутри одной колонки все данные должны быть одного типа. Проиллюстрируем это на примере:

```
> name <- c("Kolia", "Jenhia", "Sasha", "Kastia", "Vasia", " Jora")
> weight <- c(60, 68, 71, 87, 67, 93)
> height <- c(174, 168, 178, 188, 165, 172)
> size <- c("L", "S", "XL", "XXL", "S", "M")
> sex <- c("male", "female", "male", "male", "female", "male")
> data <- data.frame(name, weight, height, size, sex)
> data
  name weight height size  sex
1 Kolia   60   174   L  male
2 Jenhia   68   168   S female
3 Sasha   71   178  XL  male
4 Kastia   87   188  XXL  male
5 Vasia   67   165   S female
6 Jora    93   172   M  male
```

Очень часто бывает нужно отобразить несколько конкретных колонок. Это можно сделать разными способами (исключаем столбец name и отобразить столбец weight):

```
> data[,-1]
  weight height size  sex
1   60   174   L  male
2   68   168   S female
3   71   178  XL  male
4   87   188  XXL  male
5   67   165   S female
6   93   172   M  male
> data[,2:4]
  weight height size
1   60   174   L
```

```

2 68 168 S
3 71 178 XL
4 87 188 XXL
5 67 165 S
6 93 172 M
> data$weight
[1] 60 68 71 87 67 93

```

Отобратить из нашей таблицы только данные, относящиеся к женщинам? Вот один из способов:

```

> data[data$sex == "female",]
  name weight height size sex
2 Jenhia  68  168   S female
5 Vasia   67  165   S female

```

Или отобразить только нужные данные с помощью команды subset:

```

> data1 <- subset(data, sex == "male")
> data1
  name weight height size sex
1 Kolia   60  174   L male
3 Sasha   71  178  XL male
4 Kastia  87  188  XXL male
6 Jora    93  172   M male
> data2 <- subset(data, height >= 168 & sex == "male")
> data2
  name weight height size sex
1 Kolia   60  174   L male
3 Sasha   71  178  XL male
4 Kastia  87  188  XXL male
6 Jora    93  172   M mal

```

Более сложным случаем селекции является сортировка таблиц данных. Для сортировки таблиц данных достаточно применить команду sort(). Например, если мы хотим отсортировать наши данные сначала по полу, а потом по росту, придется выполнить следующую команду:

```

> data[order(data$sex, data$height), ]
  name weight height size sex
5 Vasia   67  165   S female
2 Jenhia  68  168   S female
6 Jora    93  172   M male
1 Kolia   60  174   L male
3 Sasha   71  178  XL male
4 Kastia  87  188  XXL male

```

ЛАБОРАТОРНАЯ РАБОТА № 2: «Визуализация данных с помощью R»

1. Цель работы

- Освоить основные распространенные типы графиков в R
- Приобрести основные навыки работы с пакетом ggplot2 в R

2. Задачи

- Углубить и закрепить знания по основным типам графикам.

- Научить строить основные типы графики в R
- Совершенствовать навыки самостоятельной работы.

3. План выполнения работы

Хотя пакет `ggplot2` - пакет к более сложным и красивым графикам для визуализации данных в R, но это не единственный способ сделать графики в R. Для быстрого изучения данных иногда полезно использовать базовые функции в R. они устанавливаются по умолчанию с R и не требуется установить дополнительные пакеты.

Эта лаб помогает вам усвоить навыки по построению различных типов графиков в R с использованием базовых функций и пакета `ggplot2`.

3.1 Scatter Plot (диаграмма рассеяния)

- Для построения Scatter Plot (Рис.1) воспользуемся `plot()`:

`plot(mtcars$wt, mtcars$mpg)`

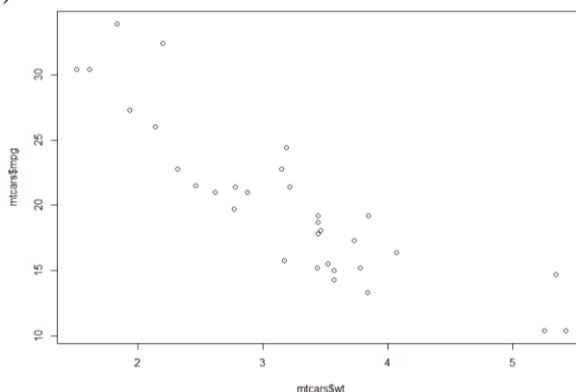


Рис 1. Scatter plot с базовой функцией `plot`

- С пакетом `ggplot2`, вы также можете получить похожий результат использованием функции `qplot()` (рис 2):

`library(ggplot2)`

`qplot(mtcars$wt, mtcars$mpg)`

- Если два вектора в одном фрейме данных, вы можете использовать следующий синтаксис:

`qplot(wt, mpg, data=mtcars)`

это эквивалентно:

`ggplot(mtcars, aes(x=wt, y=mpg)) + geom_point()`

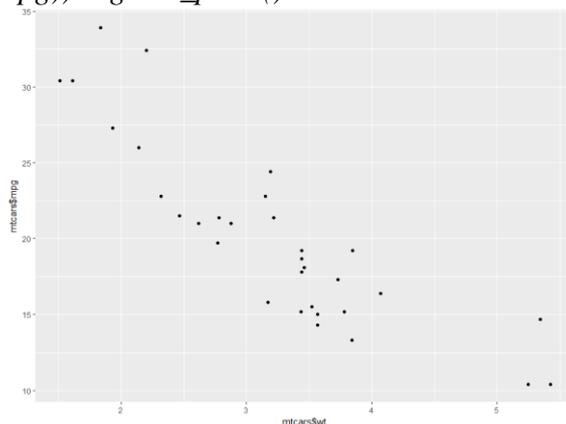


Рис 2. Scatter plot с `qplot()` из `ggplot2`

3.2 Line Graph (Линейный график)

- Для построения Line Graph воспользуемся `plot()` и указать тип графика `type = "l"` (Рис.3, слева):

`plot(pressure$temperature, pressure$pressure, type="l")`

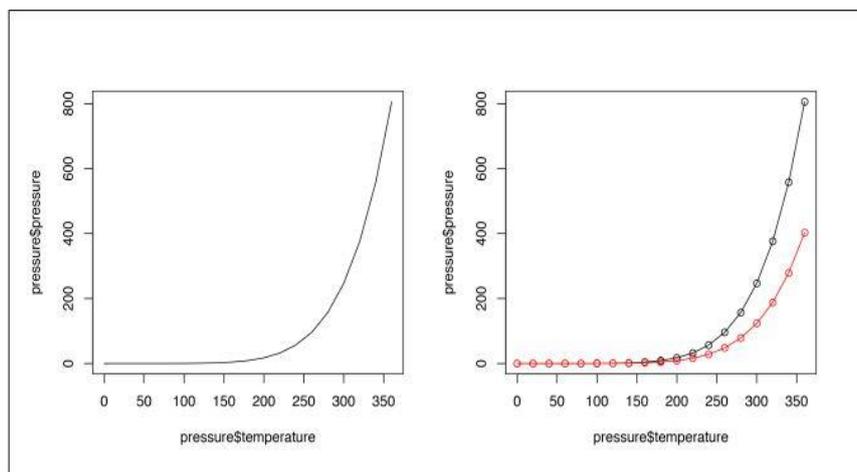


Рис 3. Line graph с базовой функцией, справа: с точками и другой линией

- Для добавления точек и/или линий (Рис.3б справа), сначала вызываем функцию `plot()` для первой линии, и потом добавляем точки с использованием `points()` и добавляем другие линии с `lines()`:

```
plot(pressure$temperature, pressure$pressure, type="l")
points(pressure$temperature, pressure$pressure)
lines(pressure$temperature, pressure$pressure/2, col="red")
points(pressure$temperature, pressure$pressure/2, col="red")
```

- с `ggplot2`, вы также можете получить похожий результат использованием функции `qplot()` с `geom = "line"` (рис. 4):

```
library(ggplot2)
qplot(pressure$temperature, pressure$pressure, geom="line")
```

- Если два вектора в одном фрейме данных, вы можете использовать следующий синтаксис:

```
qplot(temperature, pressure, data=pressure, geom="line")
# This is equivalent to:
ggplot(pressure, aes(x=temperature, y=pressure)) + geom_line()
# Lines and points together
qplot(temperature, pressure, data=pressure, geom=c("line", "point"))
# Equivalent to:
ggplot(pressure, aes(x=temperature, y=pressure)) + geom_line() + geom_point()
```

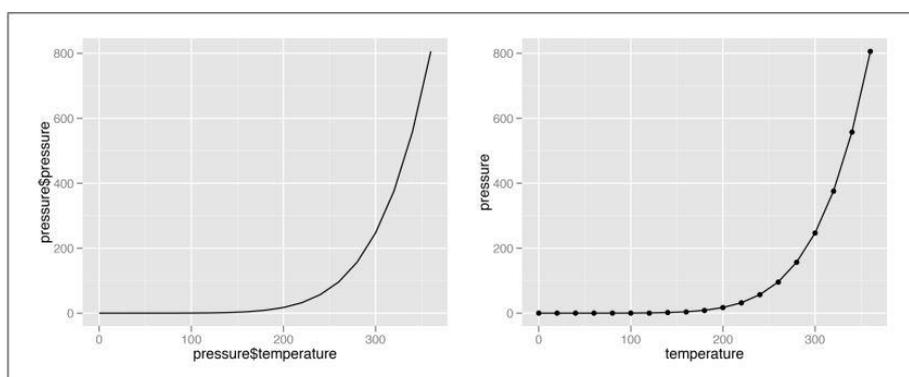


Рис 4. Line graph с `qplot()` из `ggplot2`, справа: с добавлением точек

3.3 Bar Graph (столбчатая диаграмма)

- для построения Bar Graph (рис. 5) воспользуемся `barplot()` и передаем ему вектор значений и вектор надписей:

```
barplot(BOD$demand, names.arg=BOD$Time)
```

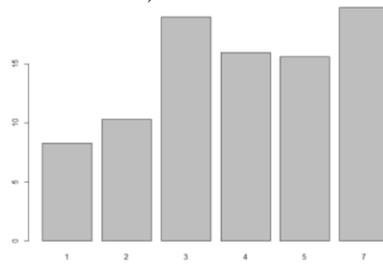


Рис 5. Bar Graph с базовой функцией barplot()

- С пакетом ggplot2 , вы также можете получить похожий результата использованием qplot(). Обратите внимание на разницу в результате, когда переменная x- непрерывная и когда она дискретная :

```
library(ggplot2)
```

```
qplot(BOD$Time, BOD$demand, geom="bar", stat="identity")
```

```
# Convert the x variable to a factor, so that it is treated as discrete
```

```
qplot(factor(BOD$Time), BOD$demand, geom="bar", stat="identity")
```

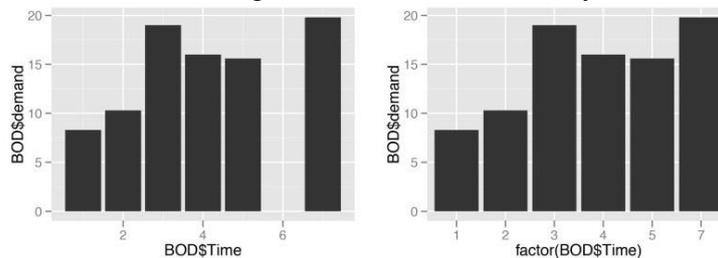


Рис 6. Bar Graph с qplot() с непрерывной переменной x, справа: с переменной, преобразующей в фактор (здесь нет надписи 6)

- Если вектор в фрейме данных, вы можете использовать следующий синтаксис:

```
qplot(Time, demand, data=BOD, geom="bar", stat="identity")
```

```
# This is equivalent to:
```

```
ggplot(BOD, aes(x=Time, y=demand)) + geom_bar(stat="identity")
```

3.4 Histogram (гистограмма)

- Для построения Histogram (рис.7) воспользуемся hist():

```
hist(mtcars$mpg)
```

```
# Specify approximate number of bins with breaks
```

```
hist(mtcars$mpg, breaks=10)
```

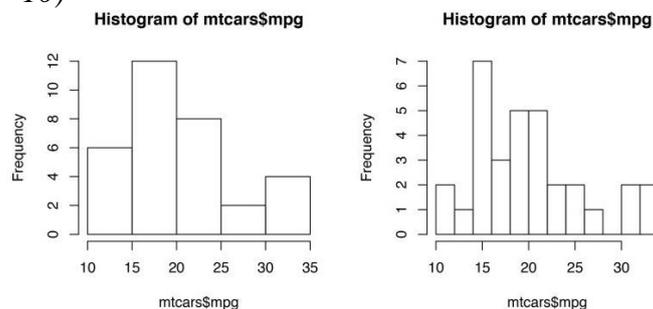


Рис 7. Histogram с базовой функцией, справа: с более ящиками

- С пакетом ggplot2 вы также можете получить похожий результат использованием qplot() (рис. 8):

```
qplot(mtcars$mpg)
```

- Если вектор в фрейме данных вы можете использовать следующий синтаксис:

```
library(ggplot2)
qplot(mpg, data=mtcars, binwidth=4)
# This is equivalent to:
ggplot(mtcars, aes(x=mpg)) + geom_histogram(binwidth=4)
```

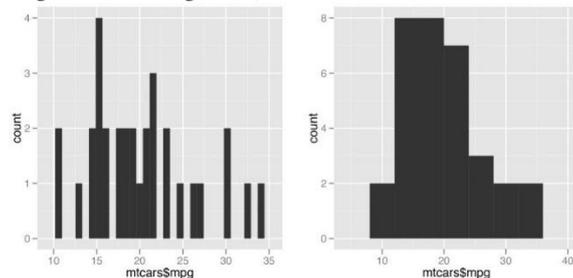


Рис 8. Histogram с qplot () из ggplot2 с шириной ящика по умолчанию, справа: с широкими ящиками.

3.5 Box Plot (ящики с усами)

- Для построения box plot (рис. 9) воспользуемся plot(). Если x – фактор, то функция plot() автоматически строит box plot:

```
plot(ToothGrowth$supp, ToothGrowth$len)
```

-Если 2 вектора в одном фрейме данных, вы можете использовать следующий синтаксис (рис. 9):

```
# Formula syntax
```

```
boxplot(len ~ supp, data = ToothGrowth)
```

```
# Put interaction of two variables on x-axis
```

```
boxplot(len ~ supp + dose, data = ToothGrowth)
```

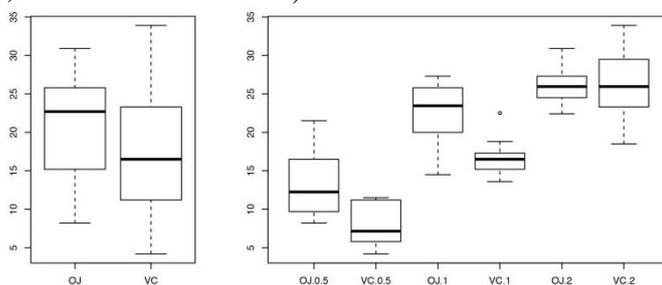


Рис 9. Слева: box plot с базовой функцией, справа: с несколькими группирующими переменными

- С пакетом ggplot2 вы также можете получить похожий результат использованием qplot() (рис. 10) с geom = “boxplot”:

```
library(ggplot2)
```

```
qplot(ToothGrowth$supp, ToothGrowth$len, geom="boxplot")
```

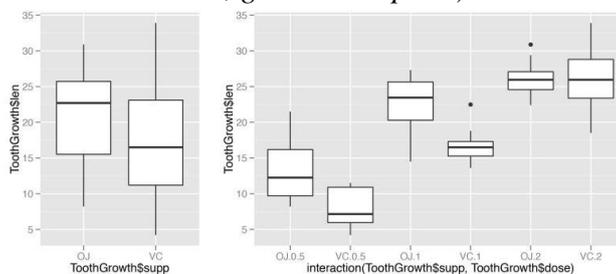


Рис 10. Слева: box plot с qplot() из ggplot2, справа: с несколькими группирующими переменными

-Если 2 вектора в одном фрейме данных, вы можете использовать следующий синтаксис:

```
qplot(supp, len, data=ToothGrowth, geom="boxplot")
# This is equivalent to:
ggplot(ToothGrowth, aes(x=supp, y=len)) + geom_boxplot()
```

ЛАБОРАТОРНАЯ РАБОТА № 3: «Прогнозирование временных рядов в R»

1. Цель работы

- Освоить основные модели прогнозирования временных рядов
- Приобрести основные навыки работы с пакетами в R для прогнозирования

2. Задачи

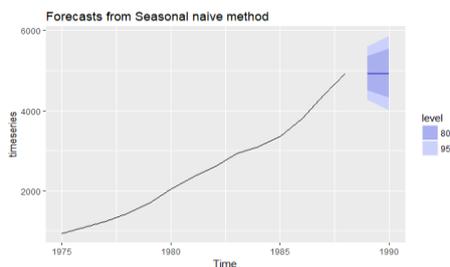
- Углубить и закрепить знания по основным моделям прогнозирования временных рядов.
- Научить использовать основные пакеты в R для прогнозирования
- Совершенствовать навыки самостоятельной работы.

3. План выполнения работы

3.1 Benchmark forecasting

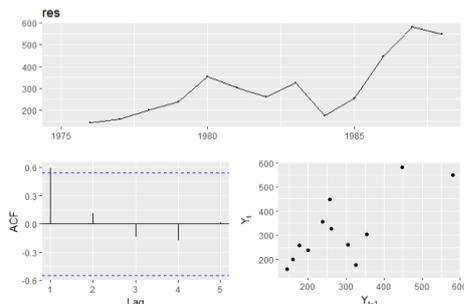
- Загрузка данных для прогнозирования (в качестве примера воспользуемся данными из M3 Competition)
 - Вызываем необходимые пакеты: ggplot2, forecasting, Mcomp
 - Построение графика данных с прогнозами путём использования наиболее распространенных Benchmark методов: mean, naïve, seasonal or drift
- Например:

```
> library(Mcomp)
> timeseries <- M3[[1]]$x
> fc <- snaive(timeseries)
> autoplot(fc)
```



- Вычислить residuals и построить их ACF:

```
> res <- residuals(fc)
> ggtsdisplay(res, plot.type="scatter")
```



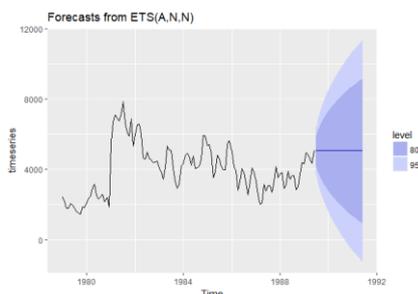
3.2 Exponential smoothing

- Загрузка данных для прогнозирования (в качестве примера воспользуемся данными из M3 Competition)
- Вызываем необходимые пакеты: ggplot2, forecasting, Mcomp

- Построение графика данных с прогнозами путём использования ETS модели из пакета forecast

Например:

```
> library(Mcomp)
> timeseries <- M3[[2000]]$x
> fitets <- ets(timeseries)
> fcfitets <- forecast(fitets)
> autoplot(fcfitets)
```



3.3 ARIMA models

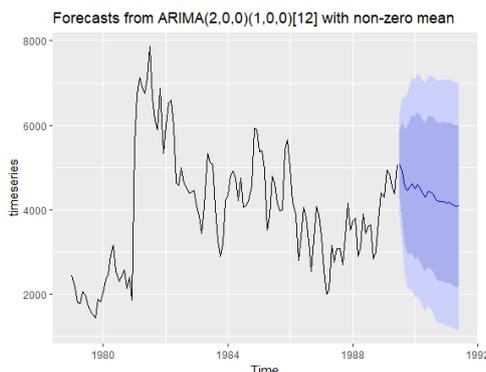
- Загрузка данных для прогнозирования (в качестве примера воспользуемся данными из M3 Competition)

- Вызываем необходимые пакеты: ggplot2, forecasting, Mcomp

- Построение графика данных с прогнозами путём использования ARIMA модели из пакета forecast

Например:

```
> timeseries <- M3[[2000]]$x
> arimats <- auto.arima(timeseries)
> fcarimats <- forecast(arimats)
> autoplot(fcarimats)
```



3.4 Оцка качества прогнозирования (forecast accuracy)

- Изучение основных метриков для оценки качества моделей прогнозирования:

- Использование R пакетов для оценки качества моделей прогнозирования: **forecast**, **FORA**

- **FORA**: R пакет, разработанный Andrey Davydenko, Maxim Shcherbakov and Sai Van Cuong из ВолГТУ для визуализации характеристик прогнозов и сравнения качества различных моделей прогнозирования с разными горизонтами прогнозирования,....

ЛАБОРАТОРНАЯ РАБОТА № 4: «Классификация и кластеризация в R»

1. Цель работы

- Освоить основные методы классификации и кластеризации
- Приобрести основные навыки работы с классификацией и кластеризацией временных рядов в R

2. Задачи

- Углубить и закрепить знания по основным методам классификации и кластеризации.
- Научить делать классификацию и в R
- Совершенствовать навыки самостоятельной работы.

3. План выполнения работы

3.1 Основные алгоритмы классификации и кластеризации

- Алгоритмы классификации:

- Linear classifiers
 - Logistic regression
 - Naive Bayes classifier
 - Fisher's linear discriminant
- Support vector machines
 - Least squares support vector machines
- Quadratic classifiers
- Kernel estimation
 - k-nearest neighbor
- Decision trees
 - Random forests
- Neural networks
- Learning vector quantization

- Алгоритмы кластеризации:

- Linear clustering algorithm
 - k-means clustering algorithm
 - Fuzzy c-means clustering algorithm
 - Hierarchical clustering algorithm
 - Gaussian(EM) clustering algorithm
 - Quality threshold clustering algorithm
- Non-linear clustering algorithm
 - MST based clustering algorithm
 - kernel k-means clustering algorithm
 - Density-based clustering algorithm

Для проверки алгоритмов классификации и кластеризации в R воспользуемся распространённым набором данных **IRIS**-классификация цветков ириса по небольшому набору параметров. **IRIS** состоит из данных о 150 экземплярах ириса, по 50 экземпляров из трёх видов — Ирис щетинистый (*Iris setosa*), Ирис виргинский (*Iris virginica*) и Ирис разноцветный (*Iris versicolor*). Для каждого экземпляра измерялись четыре характеристики (в сантиметрах):

1. Длина наружной доли околоцветника (англ. *sepal length*);
2. Ширина наружной доли околоцветника (англ. *sepal width*);
3. Длина внутренней доли околоцветника (англ. *petal length*);
4. Ширина внутренней доли околоцветника (англ. *petal width*).

3.2. На основании набора данных IRIS требуется построить правило классификации, определяющее вид растения по данным измерений.

3.3 Кластеризация набора данных IRIS

**Типовые задания контрольной работы
по дисциплине «Системы обработки больших данных»**

Уметь – ПК-1ИИП.2 У-1

Контрольная работа содержит три задания, в которых требуется с помощью языка R выполнить:

Задание 1. Пусть государство вводит потоварный налог на производителей в размере t за единицу продукции. Известно, что равновесное количество сократилось на 30 единиц. Найти равновесную цену и количество. Пусть государство вводит потоварный налог на производителей в размере t за единицу продукции. Чему равна ставка налога t ?

- Найти величину потребительского излишка в равновесии.
- Решение изобразить графически. Добавить название осей.
- Добавить легенду для каждой линии графика.
- Изменить символ точки равновесия на графике ((26 вариант = 1, 27 вариант = 2) или любой пользовательский символ).
- Обозначить область на графике потребительского излишка в равновесии.
- Добавить заголовок к графику (уравнение точки равновесия).
- Добавить подзаголовок к графику (включает свою фамилию).

Уметь – ПК-4.1ИИП. У-1, ПК-4.1ИИП. У-2, ПК-4.1ИИП. У-3

Задание 2. По данным исследования распределения доходов, в одной из стран кривая Лоренца, может быть описана уравнением $y = f(x)$, где x - доля населения, y - доля доходов населения. Вычислить коэффициент Джини, провести экономический анализ.

- Построить кривую Лоренца и установить пределы для осей от 0 до 1.
- Добавить вертикальную и горизонтальную линию в точке 1, выбрав тип линии, начиная со 2 типа. Добавить подписи к этим линиям.
- Добавить заголовок.
- Добавить подзаголовок к графику (включает свою фамилию).
- Добавить подписи осей.
- Внутри графика подписать точки пересечения.
- Добавить легенду для каждой линии графика.

Задание 3. Компьютерная фирма разрабатывает и реализует аппаратно-программные комплексы K различных типов: K_i для защиты корпоративных данных. Себестоимость C разработки комплексов составляет соответственно C_i руб. За рассматриваемый период выпущено и продано Q_i штук комплексов каждого типа, цена реализации P каждого комплекса составила P_i руб. Определить прибыль и уровень рентабельности от реализации комплексов при значении постоянных затрат F руб. - список использованных источников; - приложения (при наличии).

Варианты

Задание 1

1. На рынке компьютерных игр есть две категории покупателей – геймеры и новички. Спрос геймеров задаётся функцией $Q = 54 - 2/3P$, а спрос новичков задаётся функцией $Q = 42 - 2/3P$. Предложение на рынке компьютерных игр имеет вид $Q = -5 + P$. Пусть государство вводит потоварный налог на производителей в размере t за единицу продукции. Известно, что равновесное количество сократилось на 30 единиц

Задание 2

1) $y = 0,43x^2 + 0,57x$

2) $y = 0,56x^3 + 0,44x^2$

3) $y = 0,17x^3 + 0,83x^2$

4) $y = 0,28x^2 + 0,72x$

5) $y = 0,07x^2 + 0,93x$

6) $y = 0,06x^3 + 0,94x^2$

7) $y = 0,78x^3 + 0,22x^2$

8) $y = 0,58x^2 + 0,42x$

9) $y = 0,85x^2 + 0,15x$

10) $y = 2x - 1$

Задание 3

1) Пусть значения С, Q, P, F составляют: С = (61,5; 59,8; 56,9; 62,4; 63,3; 62,8; 62,3; 61,9; 65,1; 59,4) тыс. руб., Q = (55, 61, 65, 58, 58, 57, 57, 55, 61, 61) шт., P = (326, 326, 327, 334, 335, 336, 336, 337, 337, 338) тыс. руб., F = 1300 тыс. руб.

2) Пусть значения С, Q, P, F составляют: С = (64; 62,8; 60,4; 62,2; 60,2; 60,9; 62; 63,4; 63,8; 62,7) тыс. руб., Q = (55, 56, 61, 54, 62, 58, 54, 54, 56, 59) шт., P = (339, 340, 342, 344, 345, 345, 348, 351, 351, 351) тыс. руб., F = 1400 тыс. руб.

3) Пусть значения С, Q, P, F составляют: С = (63,3; 63,8; 61; 59,4; 58,1; 60,4; 62,5; 62,2; 60,5; 60,9) тыс. руб., Q = (56, 55, 57, 62, 62, 58, 57, 57, 61, 57) шт., P = (351, 352, 353, 353, 353, 354, 355, 357, 357, 357) тыс. руб., F = 1450 тыс. руб.

4) Пусть значения С, Q, P, F составляют: С = (60; 59,8; 60,7; 59,5; 61,9; 58,2; 64,1; 64; 60,8; 61,8) тыс. руб., Q = (57, 57, 59, 58, 58, 59, 59, 54, 59, 55) шт., P = (341, 354, 453, 354, 343, 454, 345, 354, 457, 347) тыс. руб., F = 2450 тыс. руб.

5) Пусть значения С, Q, P, F составляют: С = (60,8; 61,8; 61,2; 61,1; 65,2; 58,4; 63,1; 62,4; 61,8; 63,8) тыс. руб., Q = (59, 55, 56, 56, 56, 63, 56, 58, 55, 56) шт., P = (441, 454, 443, 454, 443, 444, 445, 454, 447, 447) тыс. руб., F = 1750 тыс. руб.

6) Пусть значения С, Q, P, F составляют: С = (63,3; 60,7; 60,9; 61,9; 60,9; 61,6; 59,3; 61; 59,3; 62,6) тыс. руб., Q = (60, 60, 61, 54, 55, 58, 62, 59, 61, 57) шт., P = (541, 554, 543, 554, 543, 544, 545, 554, 547, 547) тыс. руб., F = 1950 тыс. руб.

7) Пусть значения С, Q, P, F составляют: С = (63; 63,2; 60,9; 62,6; 62,5; 62,1; 61,5; 61,4; 62; 63,3) тыс. руб., Q = (57, 55, 57, 59, 57, 56, 59, 56, 55, 57) шт., P = (241, 254, 243, 254, 243, 244, 245, 254, 247, 227) тыс. руб., F = 1340 тыс. руб.

8) Пусть значения С, Q, P, F составляют: С = (61,8; 60,7; 61,5; 63,1; 62,9; 62,5; 63,7; 59,2; 59,9; 62,4) тыс. руб., Q = (58, 58, 60, 56, 59, 57, 57, 60, 58, 54) шт., P = (376, 376, 377, 374, 375, 376, 376, 377, 377, 378) тыс. руб., F = 1670 тыс. руб.

9) Пусть значения С, Q, P, F составляют: С = (62,8; 62,6; 63,4; 62,1; 62,9; 61,6; 57,9; 62,3; 61,2; 60,8) тыс. руб., Q = (60, 59, 59, 60, 58, 56, 60, 59, 58, 59) шт., P = (312, 323, 334, 345, 356, 367, 378, 378, 389, 390) тыс. руб., F = 1170 тыс. руб.

10) Пусть значения С, Q, P, F составляют: С = (60,7; 62,9; 62,5; 55,1; 61,6; 62,4; 63,8; 57,5; 59,4; 66,3) тыс. руб., Q = (58, 58, 57, 69, 56, 57, 56, 58, 62, 62) шт., P = (122, 233, 344, 455, 566, 677, 788, 898, 909, 290) тыс. руб., F = 3170 тыс. руб.

Типовой комплект заданий для тестов

Полный комплект тестовых материалов по входному тестированию размещен на образовательном портале «АГАСУ»

Типовые тесты для входного тестирования

1. Статистическая совокупность – это:
 - а) совокупность статистических показателей, отражающая взаимосвязи, которые объективно существуют между явлениями;
 - б) конкретные численные значения статистических показателей;
 - в) совокупность социально-экономических объектов или явлений общественной жизни, объединенных некоей качественной основой, общей связью, но отличающихся друг от друга отдельными признаками.
2. Признак – это:
 - а) изменение величины либо значения признака;
 - б) качественная особенность единицы совокупности;
 - в) первичный элемент статистической совокупности.
3. Сводкой называется:
 - а) объединение единиц совокупности в некоторые группы, имеющие свои характерные особенности, общие черты и сходные размеры изучаемого признака;
 - б) особая стадия статистического исследования, в ходе которой систематизируются первичные материалы статистического наблюдения;
 - в) объект, характеризующийся цифрами.
4. Группировка – это:
 - а) упорядочение единиц совокупности по признаку;
 - б) разбивка единиц совокупности на группы по признаку;
 - в) обобщение единичных фактов.
5. Группировка, выявляющая взаимосвязи между явлениями и их признаками, называется:
 - а) типологической;
 - б) структурной;
 - в) аналитической.
6. Группировка, в которой разнородная совокупность разбивается на однородные группы, называется:
 - а) типологической;
 - б) структурной;
 - в) аналитической.
7. Группировочным признаком при построении аналитической группировки выступает:
 - а) факторный;
 - б) результативный;
 - в) факторный и результативный.
8. Основанием группировки может быть:
 - а) качественный признак;
 - б) количественный признак;
 - в) как качественный, так и количественный признаки.

9. Охарактеризуйте вид ряда распределения абитуриентов по результатам сдачи вступительных экзаменов:

Группы абитуриентов по результатам сдачи экзаменов	Число абитуриентов	Удельный вес, % к итогу
Не поступившие	50	25
Поступившие	150	75
Итого	200	100

- а) дискретный вариационный;
 б) интервальный вариационный;
 в) атрибутивный.

10. Охарактеризуйте вид ряда распределения коммерческих банков по величине работающих активов:

Величина работающих активов банка, млн. руб.	Число банков	Удельный вес, % к итогу
До 7,0	4	13,3
7,0-12,0	5	16,7
12,0-17,0	10	33,3
17,0-22,0	6	20,0
22,0 и более	5	16,7
Итого	30	100,0

- а) дискретный вариационный;
 б) интервальный вариационный;
 в) атрибутивный.

11. Представлен макет статистической таблицы, характеризующий группировку промышленных предприятий по среднегодовой стоимости основных фондов:

Стоимость основных фондов предприятия, млн. руб.	Число предприятий	Объем выпускаемой продукции, млн. руб.		Численность промышленно-производственного персонала, чел.	
		Всего	В среднем на одно предприятие	Всего	В среднем на одно предприятие
10,0-12,0					
12,0-14,0					
14,0-16,0					
Итого					

Какой вид группировки отражает данный макет:

- а) типологическую;
 б) структурную;
 в) аналитическую.

12. Распределение предприятий по тарифному разряду характеризуется следующими данными:

Тарифный разряд	Число рабочих	Удельный вес, % к итогу
-----------------	---------------	-------------------------

2	5	10
3	6	12
4	5	10
5	12	24
6	22	44
Итого	50	100

Определите вид ряда распределения: а) интервальный вариационный;

б) дискретный вариационный;

в) атрибутивный.

13. Абсолютные статистические показатели выражаются:

а) в процентах;

б) в именованных числах;

в) в коэффициентах.

14. Показатели, выражающие размеры, объем, уровни социально-экономических явлений и процессов, являются величинами:

а) абсолютными;

б) относительными.

15. В каких единицах измерения могут выражаться абсолютные величины:

а) натуральных;

б) трудовых и денежных;

в) условно-натуральных;

г) отвлеченных.

16. Относительные величины выполнения плана исчисляются как:

а) отношение планового задания на предстоящий период к фактически достигнутому уровню, являющемуся базисным для плана;

б) отношение фактически достигнутого уровня к плановому заданию за тот же период времени.

17. Относительные величины динамики получаются в результате сопоставления показателей каждого последующего периода:

а) с предыдущим;

б) с первоначальным;

в) со средним.

18. Относительные величины структуры:

а) характеризуют состав явления и показывают, какой удельный вес в общем итоге составляет каждая его часть;

б) показывают соотношение отдельных составных частей целого явления.

19. Относительные величины интенсивности представляют собой:

а) отношение двух разноименных показателей, находящихся в определенной взаимосвязи;

б) отношение двух одноименных показателей, относящихся к разным объектам или территориям за один и тот же период или момент времени.

Типовой комплект заданий для тестов

Полный комплект тестовых материалов по итоговому тестированию размещен на образовательном портале «АГАСУ»

Типовые тесты для итогового тестирования

Знать – ПК-1ИИП.2 У-1

1. Для измерения вариации значения признака не вычисляют показатели: а) моду; б) дисперсию; в) размах вариации; г) среднелинейное отклонение; д) коэффициент вариации.
2. Модой в статистике называют:
а) значение признака, которое чаще всего встречается в данной совокупности;
б) значение признака у единицы, которое находится в середине упорядоченного ряда распределения;
в) значение признака, которое встречается в данной совокупности единственный раз.
3. Если данные сгруппированы, но каждое значение признака встречается неодинаковое число раз, то для осреднения признака применяется формула (запишите ее):
а) средняя гармоническая простая;
б) средняя хронологическая простая;
в) средняя арифметическая взвешенная;
г) средняя гармоническая взвешенная.
4. Если известны значения признака у каждой единицы и количество единиц, обладающих тем или иным значением признака, то применяется формула (запишите ее):
а) средняя хронологическая;
б) средняя арифметическая простая;
в) средняя арифметическая взвешенная.
5. Средний уровень интервального ряда динамики с равноотстоящими уровнями определяется по формуле:
а) средней гармонической;
б) средней арифметической простой;
в) средней хронологической.
6. Средний уровень моментного ряда динамики определяется по формуле:
а) средней гармонической взвешенной;
б) средней хронологической;
в) средней арифметической простой.
7. Прием обнаружения общей тенденции развития не являются:
а) метод скользящей средней;
б) аналитическое выравнивание ряда динамики;
в) приведение рядов динамики к одному основанию;
г) укрепление интервалов.
8. Темп роста вычисляется как:
а) отношение уровней ряда;
б) разность уровней ряда;

- в) произведение уровней ряда;
- г) разность темпа прироста и 100 %.

9. В таблицах приведены примеры рядов динамики:

Ряд динамики № 1. Объем продаж рекламного времени радиостанцией за 6 недель:

Показатель	Текущий номер недели					
	1	2	3	4	5	6
Проданное рекламное время, мин.						

Ряд динамики № 2. Среднемесячная заработная плата рабочих предприятия (руб.):

Показатель	Месяц					
	01.05	02.05	03.05	04.05	05.05	06.05
Средняя заработная плата рабочих, руб.	9570	10900	11950	1200	13100	16000

Ряд динамики № 3. Цены акций промышленной компании на момент открытия торгов (долл.):

Показатель	Дата					
	06.09.05	07.09.05	08.09.05	09.09.05	10.09.05	13.09.05
Средняя заработная плата рабочих, руб.	280	291	287	289	294	286

- а) ряд динамики № 1;
- б) ряд динамики № 2;
- в) ряд динамики № 3;
- г) пример интервального ряда динамики отсутствует.

10. Каковы отличительные особенности статистического наблюдения?

- а) научная организация,
- б) обширная программа.

11. Какие организационные формы статистического наблюдения Вам известны?

- а) анкетное;
- б) документальное;
- в) отчетность;
- г) периодическое.

12. Перечислите виды статистической отчетности.

- а) текущая;
- б) годовая;
- в) общая;
- г) комбинированная.

13. Назовите виды статистического наблюдения по охвату единиц объекта.

- а) опрос;
- б) саморегистрация;

- в) несплошное;
- г) текущее.

14. Почему переписи населения проводятся в зимнее время года?

- а) достигается экономия ресурсов на подготовку персонала;
- б) меньшие транспортные расходы;
- в) наименьшая подвижность опрашиваемых;
- г) не нужен критический момент наблюдения.

15. Укажите основной фактор, определяющий период (срок) статистического наблюдения.

- а) степень разработанности организационного плана;
- б) размеры объекта наблюдения;
- в) наличие и квалификация персонала статистического наблюдения;
- г) организационная форма.

16. На склад коммерческой организации поступила партия товара. Для проверки его качества была отобрана десятая часть партии и путем тщательного осмотра каждой единицы товара определялось и фиксировалось качество. К какому виду наблюдения по полноте охвата объекта можно отнести это обследование?

- а) монографическое;
- б) выборочное;
- в) метод основного массива.

17. Редакция журнала разослала читателям вопросник с просьбой ответить на содержащиеся в нем вопросы и вернуть в редакцию. Как называется использованный редакцией способ собирания сведений?

- а) анкетный;
- б) экспедиционный;
- в) явочный;
- г) саморегистрация.

18. Назовите виды ошибок статистического наблюдения.

- а) логические;
- б) систематические;
- в) арифметические;
- г) комбинированные.

19. Какой вид контроля точности материалов переписи населения применяется при сопоставлении данных о годе рождения и возрасте опрашиваемых?

- а) арифметический;
- б) логический;
- в) другой вид контроля;
- г) в данном случае контроль не нужен.

20. К каким группировочным признакам относятся: национальность, форма собственности, профессия рабочего?

- а) к количественным,
- б) к атрибутивным.

21. Какими группировочными признаками являются: доход сотрудника фирмы, стоимость реализованной продукции, объем произведённой продукции?

- а) количественными;
- б) атрибутивными.

22. Какими группировочными признаками являются: ликвидные активы, уставной капитал, стоимость основных фондов?

- а) дискретными;
- б) непрерывными.

23. Охарактеризуйте вид группировочного признака: тарифный разряд, балл успеваемости, число детей в семье?

- а) дискретный;
- б) интервальный.

Знать - ПК-4.1ИИП. 3-1

24. Охарактеризуйте вид ряда распределения абитуриентов по результатам сдачи вступительных экзаменов на I курс ВУЗа:

Группы абитуриентов по результатам сдачи экзаменов	Число абитуриентов	Удельный вес абитуриентов в %% к итогу
не поступившие	50	25
поступившие	150	75
Итого	200	100

- а) дискретный вариационный;
- б) интервальный вариационный;
- в) атрибутивный.

25. Объект А по величине исследуемого показателя превышает объект Б на 20 %. На сколько объект Б меньше объекта А?

- а) менее, чем на 20%;
- б) на 20%;
- в) более, чем на 20%.

26. Бывает ли относительный показатель сравнения быть именованной величиной?

- а) может, если исходные абсолютные показатели выражены в условно-натуральных единицах измерения;
- б) не может.

27. Для изображения взаимосвязи между факторным и результативным признаками на графике применяются диаграммы:

- а) столбиковые;
- б) линейные.

28. Какие виды диаграмм можно использовать по форме геометрического образа:

- а) плоскостные;
- б) статистические карты.

29. Какие виды статистических графиков существуют по способу построения:

- а) диаграммы;
- б) плоскостные.

30. Какие виды статистических графиков могут использоваться для изображения рядов распределения:

- а) диаграммы сравнения;
- б) кумуляты;
- в) диаграммы структуры.

31. Выберите способ графического изображения данных о распределении научных работников по отраслям наук на конец года по региону:
- а) картограмма;
 - б) столбиковая.
32. При изображении структуры и структурных сдвигов в совокупности явлений на графике применяются диаграммы:
- а) полосовые;
 - б) секторные.
33. При построении линейных диаграмм используются масштабные шкалы:
- а) равномерные;
 - б) логарифмические.
34. При изображении на графике сезонных колебаний применяются диаграммы:
- а) линейные;
 - б) радиальные.
35. При изображении социально-экономических явлений при помощи диаграмм применяются их виды:
- а) фоновые;
 - б) знаков-символов.
36. Изобразительными знаками в картодиаграммах являются:
- а) точки и штриховка;
 - б) круги, квадраты и другие фигуры.
37. К какому виду относится структура себестоимости продукции предприятия?
- а) моментальная;
 - б) интервальная.
38. К какому виду относится структура родившихся по полу?
- а) моментальная;
 - б) интервальная.
39. В чём измеряется «абсолютные» показатели структурных сдвигов?
- а) в процентах;
 - б) в процентных пунктах;
 - в) в процентных пунктах и в промилле.
40. В чём измеряются относительные показатели структурных сдвигов?
- а) в коэффициентах;
 - б) в процентах;
 - в) в коэффициентах и процентах.

Знать - ПК-4.1ИИП. 3-2

41. Средний «абсолютный» прирост удельного веса позволяет получить обобщающую характеристику изменения:
- а) одной структурной части;
 - б) всех структурных частей.
42. Может ли средний темп роста удельного веса быть отрицательной величиной?
- а) может;
 - б) не может.
43. Чему равна сумма средних удельных весов всех структурных частей совокупности?
- а) нулю;

- б) 100%;
- в) произвольной величине.
44. Могут ли «абсолютные» и относительные показатели структурных сдвигов приводить к противоречивым выводам?
- а) да;
- б) нет.
45. Можно ли при расчёте коэффициента Джинни использовать неравные группы?
- а) можно;
- б) нельзя.
46. В какой мере величина обобщающего показателя централизации зависит от объёма совокупности?
- а) не зависит;
- б) мало зависит;
- в) сильно зависит.
47. Возможна ли многовариантность значений среднего показателя, рассчитанного по одним и тем же данным?
- а) да;
- б) нет.
48. Могут ли взвешенные и невзвешенные средние, рассчитанные по одним и тем же данным, совпадать?
- а) да;
- б) нет.
49. Может ли одно и то же исходное соотношение быть реализовано на \bullet скове различных форм средней?
- а) да;
- б) нет.
50. Как изменится средняя величина, если все варианты признака уменьшились в 1,5 раза, а все веса в 1,5 раза увеличить?
- а) не изменится;
- б) уменьшится;
- в) возрастет.
51. Изменится ли средняя величина, если все веса уменьшить на 20%?
- а) изменится;
- б) не изменится.
52. Изменится ли средняя величина, если все веса уменьшить на некоторую постоянную величину?
- а) изменится;
- б) не изменится.
53. Могут ли мода, медиана и средняя арифметическая совпадать?
- а) могут;
- б) не могут.
54. Может ли ряд распределения характеризоваться двумя и более видами?
- а) нет;
- б) может двумя;
- в) может двумя и более.
55. В каких границах изменяется коэффициент вариации?

- а) от 0 до 100%;
б) от 0 до 200%;
в) нижняя граница — 0%, верхняя — практически отсутствует.
56. Всегда ли абсолютные и относительные показатели вариации приводят к непротиворечивым выводам?
а) всегда;
б) не всегда.
57. Сколько способов отбора получили наибольшее распространение в практике выборочных обследований?
а) менее шести;
б) шесть;
в) более шести.
58. При каком отборе одни и те же единицы подвергаются обследованию по расширяющейся программе?
а) при комбинированном;
б) при многоступенчатом;
в) при многофазном.
59. Для каких способов формирования выборочной совокупности необходимый объём выборки определяется по одинаковым формулам?
а) для собственно-случайного и механического;
б) для собственно-случайного и типического;
в) для механического и типического;
г) для других способов отбора.
60. При каком способе распространения данных выборочного наблюдения на генеральную совокупность используется понятие поправки на недоучёт?
а) при способе прямого пересчёта;
б) при способе коэффициентов.
61. При какой минимальной численности единиц выборочной совокупности используются формулы малой выборки?
а) 10;
б) 20;
в) 30;
г) 50;
д) 100.
62. Доверительный интервал выборочной средней и доли при малой выборке является односторонним или двусторонним?
а) односторонним;
б) двусторонним.
63. Каким должно быть распределение признака в генеральной совокупности, чтобы результаты малой выборки имели практическое значение?
а) нормальным;
б) асимптотически нормальным;
в) нормальным и асимптотически нормальным.

**Лист внесения дополнений и изменений
в рабочую программу учебной дисциплины
«Системы обработки больших данных»
(наименование дисциплины)**

на 2025 - 2026 учебный год

Рабочая программа пересмотрена на заседании кафедры «САПРиМ»,
протокол № 9 от 14 апреля г.

Зав. кафедрой

к.пед.н

ученая степень, ученое звание



(подпись)

/ В.В. Соболева /

И.О. Фамилия

В рабочую программу вносятся следующие дополнения:

1. В рабочую программу вносятся следующие изменения:

1. П.8.1 представлен в следующей редакции:

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

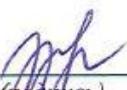
а) основная учебная литература:

1. Татарникова, Т. М. Интеллектуальный анализ данных : учебное пособие / Т. М. Татарникова. — Москва, Вологда : Инфра-Инженерия, 2024. — 172 с. — ISBN 978-5-9729-1772-3. — Текст : электронный // Цифровой образовательный ресурс IPR SMART : [сайт]. — URL: <https://www.iprbookshop.ru/143351.html>
2. Алексеев, В. Е. Структуры данных и модели вычислений : учебное пособие / В. Е. Алексеев, В. А. Таланов. — 4-е изд. — Москва : Интернет-Университет Информационных Технологий (ИНТУИТ), Ай Пи Ар Медиа, 2025. — 247 с. — ISBN 978-5-4497-0939-4. — Текст : электронный // Цифровой образовательный ресурс IPR SMART : [сайт]. — URL: <https://www.iprbookshop.ru/146401.html>

Составители изменений и дополнений:

к.т.н., доц.

ученая степень, ученое звание



(подпись)

/ С.В.Окладникова /

И.О. Фамилия

Председатель МКС «Информационные системы и технологии (профиль) «Искусственный интеллект в проектировании городской среды»

д.т.н., проф.

ученая степень, ученое звание



(подпись)

/ В.В. Соболева /

И.О. Фамилия