

Документ подписан простой электронной подписью

Информация о владельце:

ФИО: Волков В.В.

Должность: Ректор

Дата подписания: 14.08.2025 17:03:05

Уникальный программный ключ:

ed68fd4b85b778e0f0b1bfea5dbc56cf4148f1225917e739a70e315174f88591

**Автономная некоммерческая образовательная организация высшего образования
«Европейский университет в Санкт-Петербурге»**

Школа вычислительных социальных наук

УТВЕРЖДАЮ:

Ректор

В.В. Волков

« 26 »

августа

2025 г.

Протокол УС № 2

от 26.02.2025 г.



**Рабочая программа дисциплины
Введение в программирование на R**

**образовательная программа
направление подготовки
39.04.01 Социология**

**направленность (профиль) программы
«Вычислительная социология»
уровень высшего образования – магистратура**

Программа двух квалификаций:

- «магистр» по направлению подготовки 39.04.01 Социология;
- дополнительная квалификация – «магистр» по направлению подготовки 09.04.03 Прикладная информатика

язык обучения – русский
форма обучения - очная

Санкт-Петербург

Автор:

Аркатов Д. А., к.п.н, доцент Школы вычислительных социальных наук по направлению Социология АНООВО «ЕУСПб»

Рецензент: Тенишева К.А., кандидат социологических наук, доцент направления Социология, директор программ по направлению Социология Школы Вычислительных социальных наук

Рабочая программа дисциплины **«Введение в программирование на R»**, входящей в образовательную программу уровня магистратуры «Вычислительная социология», утверждена на заседании Совета Школы вычислительных социальных наук

Протокол заседания № 04 от 25.02.2025 года

АННОТАЦИЯ РАБОЧЕЙ ПРОГРАММЫ ДИСЦИПЛИНЫ

Дисциплина **«Введение в программирование на R»** является дисциплиной базовой части Блока 1 «Дисциплины (модули)» образовательной программы «Вычислительная социология».

Дисциплина **«Введение в программирование на R»** знакомит студентов с языком программирования R и средой разработки RStudio. Учащиеся освоят основы синтаксиса, работу с данными, создание визуализаций с помощью ggplot2, а также научатся оформлять аналитические отчёты в RMarkdown. Ключевые темы: Основы программирования в R (переменные, структуры данных, функции); импорт и обработка данных (dplyr, tidyr); визуализация данных (ggplot2); создание динамических отчётов в RMarkdown. Студентам также покажут наиболее современные ИИ-инструменты, помогающие изучать программирование и упрощающие сам процесс кодинга.

Общая трудоемкость освоения дисциплины составляет 2 зачетных единицы, 72 часа.

Содержание

1. НАИМЕНОВАНИЕ, ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ	5
2. ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОСВОЕНИЯ ДИСЦИПЛИНЫ	5
3. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ.....	7
4. ОБЪЕМ ДИСЦИПЛИНЫ.....	7
5. СОДЕРЖАНИЕ И СТРУКТУРА ДИСЦИПЛИНЫ	7
5.1 Содержание дисциплины	7
5.2 Структура дисциплины.....	9
6. УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ САМОСТОЯТЕЛЬНОЙ РАБОТЫ ОБУЧАЮЩИХСЯ ПО ДИСЦИПЛИНЕ.....	9
6.1 Общие положения	9
6.2 Рекомендации по распределению учебного времени по видам самостоятельной работы и разделам дисциплины	9
6.3 Перечень основных вопросов по изучаемым темам для самостоятельной работы обучающихся по дисциплине	9
6.4 Перечень литературы для самостоятельной работы обучающегося:	12
6.5 Перечень учебно-методического обеспечения для самостоятельной работы.....	12
7. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ДЛЯ ПРОВЕДЕНИЯ ТЕКУЩЕЙ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ОБУЧАЮЩИХСЯ ПО ДИСЦИПЛИНЕ	12
7.1 Показатели, критерии и оценивание компетенций и индикаторов их достижения в процессе текущей аттестации.....	12
7.2 Контрольные задания для текущей аттестации.....	13
7.3 Показатели, критерии и оценивание компетенций и индикаторов их достижения в процессе промежуточной аттестации	13
7.4 Типовые задания к промежуточной аттестации.....	16
7.5 Средства оценки индикаторов достижения компетенций.....	19
8. ОСНОВНАЯ И ДОПОЛНИТЕЛЬНАЯ ЛИТЕРАТУРА.....	20
9. ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ, ИСПОЛЬЗУЕМЫЕ ПРИ ОСУЩЕСТВЛЕНИИ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА.....	20
9.1 Программное обеспечение	21
9.2 Перечень информационно-справочных систем и профессиональных баз данных информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины:	21
9.3 Лицензионные электронные ресурсы библиотеки Университета	21
9.4 Электронная информационно-образовательная среда Университета.....	22
10. МАТЕРИАЛЬНО-ТЕХНИЧЕСКАЯ БАЗА, НЕОБХОДИМАЯ ДЛЯ ОСУЩЕСТВЛЕНИЯ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА.....	22
ПРИЛОЖЕНИЕ 1	24

1. НАИМЕНОВАНИЕ, ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Курс направлен на формирование у студентов прочной основы в программировании на языке R как инструменте анализа и визуализации данных. Студенты изучат ключевые синтаксические конструкции языка, научатся обрабатывать, структурировать и визуализировать данные, используя как базовые средства языка, так и популярные библиотеки, включая ggplot2 и plotly. Эти навыки позволят им уверенно использовать R в дальнейших курсах по прикладной статистике, анализу данных и машинному обучению.

Задачи:

- Освоить работу в среде R и RStudio, включая установку, навигацию по интерфейсу, использование скриптов, консоли, панели переменных, графического окна, а также управление пакетами и настройку рабочего пространства.
- Понять основные типы данных и структуры в R, включая векторы, списки, матрицы, массивы, факторы и фреймы данных, а также научиться выполнять с ними базовые арифметические и логические операции, индексацию и преобразование типов.
- Развить навыки импорта, экспорта и предварительной обработки данных, включая загрузку данных из файлов разных форматов (CSV, Excel, TXT), работу с пропущенными значениями, фильтрацию, переименование переменных и преобразование структуры данных.
- Сформировать алгоритмическое мышление, через изучение условных операторов и циклов, а также освоение функционального подхода.
- Научиться разрабатывать пользовательские функции, с пониманием аргументов, возвращаемых значений, области видимости переменных и построения модульного кода, пригодного для повторного использования.
- Освоить основы визуализации данных, используя как встроенные средства графики (base R), так и современные библиотеки ggplot2 для создания статических графиков и plotly для интерактивной визуализации; научиться подбирать тип визуализации в зависимости от типа данных и задачи.
- Понять принципы reproducible анализа, включая написание чистого, читаемого и воспроизводимого кода, использование проектов в RStudio, базовое документирование, создание отчетов с помощью R Markdown с включением кода, графиков и результатов анализа.

2. ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОСВОЕНИЯ ДИСЦИПЛИНЫ

В результате изучения учебной дисциплины обучающийся должен овладеть следующими компетенциями: общепрофессиональными (ОПК). Планируемые результаты формирования компетенций и индикаторы их достижения в результате освоения дисциплины представлены в Таблице 1.

Таблица 1

Планируемые результаты освоения дисциплины, соотнесенные с индикаторами достижения компетенций обучающихся

Код и наименование компетенции	Индикаторы достижения компетенции	Результаты обучения (знать, уметь, владеть)
ОПК-1 (ПИ) Способен самостоятельно приобретать, развивать и применять математические,	ИД.ОПК-1.1. Решает нестандартные задачи профессиональной деятельности с применением математических, естественнонаучных, социально-экономических методов	Знать: математические, естественнонаучные и социально-экономические методы для использования в профессиональной деятельности

Код и наименование компетенции	Индикаторы достижения компетенции	Результаты обучения (знать, уметь, владеть)
естественнонаучные, социально-экономические и профессиональные знания для решения нестандартных задач, в том числе в новой или незнакомой среде и в междисциплинарном контексте	ИД.ОПК-1.2. В условиях неопределенности и риска способен выработать эффективную стратегию действий, используя математические, естественнонаучные, социально-экономические методы науки	З (ОПК-1)
		Уметь: решать нестандартные профессиональные задачи, в том числе в новой или незнакомой среде и в междисциплинарном контексте, с применением математических, естественнонаучных социально-экономических и профессиональных знаний У (ОПК-1)
		Владеть: навыком применять математические, естественнонаучные, социально-экономические и профессиональные знания для решения нестандартных задач В (ОПК-1)
ОПК-4 (ПИ) Способен применять на практике новые научные принципы и методы исследований	ИД.ОПК-4.1. На основе современных теорий и концепций обосновывает актуальность постановки целей и задач научных исследований в профессиональной области знаний ИД.ОПК-4.2. Анализирует новые научные принципы и методы исследований в профессиональной области знаний ИД.ОПК-4.3. Применяет новые научные принципы и методы исследований в профессиональной области знаний ИД.ОПК-4.4. Разрабатывает предложения и рекомендации по использованию новых научных принципов и методов исследований в профессиональной области знаний	Знать: актуальные направления применения новых научных принципов и методов исследований в профессиональной деятельности З (ОПК-4)
		Уметь: самостоятельно формировать планы и программы научных исследований с применением новых принципов и методов, характерных для выбранной отрасли науки У (ОПК-4)
		Владеть: навыками системного использования различных новых научных принципов и методов исследований для различных направлений науки В (ОПК-4)

В результате освоения дисциплины магистрант должен:

- знать:

- основные типы данных и структуры в языке R: векторы, списки, матрицы, фреймы данных, факторы;
- синтаксис языка программирования R, включая условные конструкции, циклы, функции и принципы модульного программирования;
- подходы к визуализации данных, включая статические и интерактивные графики;
- этапы работы с данными: импорт, очистка, трансформация, подготовка к анализу.

- уметь:

- писать корректный и читаемый код на языке R для базовой обработки и анализа данных;
- применять условные конструкции, циклы и пользовательские функции при решении прикладных задач;
- визуализировать данные с использованием различных средств визуализации;
- создавать воспроизводимые аналитические отчеты, включающие код, графики и интерпретацию результатов.

- владеть:

- практическими навыками программирования на языке R в контексте анализа данных;
- методами структурирования, фильтрации и преобразования данных в табличном формате;
- инструментами визуального анализа и представления данных в исследовательской и прикладной работе;
- основами построения воспроизводимых исследовательских проектов и отчетов в научной и учебной деятельности.

3. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ

Дисциплина «Введение в программирование на R» является дисциплиной обязательной части Блока 1 «Дисциплины (модули)» образовательной программы «Вычислительная социология». Курс читается в первом модуле, форма промежуточной аттестации – зачет с оценкой.

Знания, умения и навыки, полученные при освоении данной дисциплины, применяются магистрантами в процессе прохождения учебной и производственной практики, выполнения выпускной квалификационной работы.

4. ОБЪЕМ ДИСЦИПЛИНЫ

Общая трудоемкость освоения дисциплины составляет 2 зачетных единиц, 72 часа.

Таблица 2

Объем дисциплины

Типы учебных занятий и самостоятельная работа		Объем дисциплины										
		Всего	Объем дисциплины									
			Модуль									
			1	2	3	4	5	6	7	8	9	10
Контактная работа обучающихся с преподавателем в соответствии с УП:		28	28	-	-	-	-	-	-	-	-	-
Лекции (Л)		14	14	-	-	-	-	-	-	-	-	-
Практические занятия (ПЗ)		14	14	-	-	-	-	-	-	-	-	-
Самостоятельная работа (СР)		44	44	-	-	-	-	-	-	-	-	-
Промежуточная аттестация	форм а	Зачет с оценкой	Зачет с оценкой	-	-	-	-	-	-	-	-	-
	час.	-	-	-	-	-	-	-	-	-	-	-
Общая трудоемкость дисциплины (час./з.е.)		72/2	72/2	-	-	-	-	-	-	-	-	-

5. СОДЕРЖАНИЕ И СТРУКТУРА ДИСЦИПЛИНЫ

Содержание дисциплины соотносится с планируемыми результатами обучения по дисциплине: через задачи, формируемые компетенции и их компоненты (знания, умения, навыки – далее ЗУВ) по средствам индикаторов достижения компетенций в соответствии с Таблицей 3.

5.1 Содержание дисциплины

Таблица 3

Содержание дисциплины

№ п/п	Наименование тем (разделов)	Содержание тем (разделов)	Коды компетенций	Индикаторы компетенций (в соот.с Таблицей 1)	Коды ЗУВ (в соот.с Таблицей 1)
1	Введение в R и RStudio	Установка и настройка R и RStudio. Интерфейс рабочей среды. Основные типы данных:	ОПК -1 (ПИ) ОПК -4 (ПИ)	ИД.ОПК-1.1. ИД.ОПК-1.2. ИД.ОПК-4.1. ИД.ОПК-4.2.	З (ОПК-1) У (ОПК-1) В (ОПК-1) З (ОПК-4)

№ п/п	Наименование тем (разделов)	Содержание тем (разделов)	Коды компетенций	Индикаторы компетенций (в соот.с Таблицей 1)	Коды ЗУВ (в соот.с Таблицей 1)
		числовые, логические, строковые. Выполнение простых выражений. Работа с переменными и рабочим пространством.		ИД.ОПК-4.3. ИД.ОПК-4.4.	У (ОПК-4) В (ОПК-4)
2	Основные структуры данных	Векторы, списки, матрицы, массивы, факторы. Создание, индексация, типы данных, базовые операции. Отличия и применения разных структур.	ОПК -1 (ПИ) ОПК -4 (ПИ)	ИД.ОПК-1.1. ИД.ОПК-1.2. ИД.ОПК-4.1. ИД.ОПК-4.2. ИД.ОПК-4.3. ИД.ОПК-4.4.	З (ОПК-1) У (ОПК-1) В (ОПК-1) З (ОПК-4) У (ОПК-4) В (ОПК-4)
3	Работа с фреймами данных и импортом	Импорт данных из CSV, Excel, TXT. Структура фрейма данных, доступ к строкам и столбцам, фильтрация, сортировка, создание новых переменных.	ОПК -1 (ПИ) ОПК -4 (ПИ)	ИД.ОПК-1.1. ИД.ОПК-1.2. ИД.ОПК-4.1. ИД.ОПК-4.2. ИД.ОПК-4.3. ИД.ОПК-4.4.	З (ОПК-1) У (ОПК-1) В (ОПК-1) З (ОПК-4) У (ОПК-4) В (ОПК-4)
4	Условные конструкции и циклы	Условные операторы (if, else, ifelse). Циклы (for, while, repeat). Основы функционального подхода: применение функций ко множествам элементов.	ОПК -1 (ПИ) ОПК -4 (ПИ)	ИД.ОПК-1.1. ИД.ОПК-1.2. ИД.ОПК-4.1. ИД.ОПК-4.2. ИД.ОПК-4.3. ИД.ОПК-4.4.	З (ОПК-1) У (ОПК-1) В (ОПК-1) З (ОПК-4) У (ОПК-4) В (ОПК-4)
5	Функции и модульность кода	Создание пользовательских функций. Параметры, возвращаемые значения. Структура функции. Глобальные и локальные переменные. Чистый код.	ОПК -1 (ПИ) ОПК -4 (ПИ)	ИД.ОПК-1.1. ИД.ОПК-1.2. ИД.ОПК-4.1. ИД.ОПК-4.2. ИД.ОПК-4.3. ИД.ОПК-4.4.	З (ОПК-1) У (ОПК-1) В (ОПК-1) З (ОПК-4) У (ОПК-4) В (ОПК-4)
6	Визуализация данных	Визуализация средствами base R. Построение графиков с ggplot2: точечные, линейные, гистограммы, boxplot и др. Интерактивные графики с использованием plotly.	ОПК -1 (ПИ) ОПК -4 (ПИ)	ИД.ОПК-1.1. ИД.ОПК-1.2. ИД.ОПК-4.1. ИД.ОПК-4.2. ИД.ОПК-4.3. ИД.ОПК-4.4.	З (ОПК-1) У (ОПК-1) В (ОПК-1) З (ОПК-4) У (ОПК-4) В (ОПК-4)
7	Очистка и подготовка данных к анализу	Обработка пропущенных значений. Преобразование структуры данных. Использование функций фильтрации, группировки, объединения. Подготовка к статистическому анализу.	ОПК -1 (ПИ) ОПК -4 (ПИ)	ИД.ОПК-1.1. ИД.ОПК-1.2. ИД.ОПК-4.1. ИД.ОПК-4.2. ИД.ОПК-4.3. ИД.ОПК-4.4.	З (ОПК-1) У (ОПК-1) В (ОПК-1) З (ОПК-4) У (ОПК-4) В (ОПК-4)

5.2 Структура дисциплины

Таблица 4

Структура дисциплины							
№ п/п	Наименование тем (разделов)	Объем дисциплины, час.					Форма текущего контроля успеваемости*, промежуточной аттестации
		Всего	Контактная работа обучающихся с преподавателем по типам учебных занятий в соответствии с УП			СР	
			Л	СЗ	СРП		
Очная форма обучения							
Тема 1	Введение в R и RStudio	12	2	2		8	ДЗ
Тема 2	Основные структуры данных	10	2	2		6	ДЗ
Тема 3	Работа с фреймами данных и импортом	10	2	2		6	ДЗ
Тема 4	Условные конструкции и циклы	10	2	2		6	ДЗ
Тема 5	Функции и модульность кода	10	2	2		6	ДЗ
Тема 6	Визуализация данных	10	2	2		6	ДЗ
Тема 7	Очистка и подготовка данных к анализу	10	2	2		6	ДЗ
Промежуточная аттестация		-	-	-	-	-	Зачет с оценкой
Всего:		72/2	14	14	-	44	-

*Примечание: формы текущего контроля успеваемости: домашнее задание (ДЗ).

6. УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ САМОСТОЯТЕЛЬНОЙ РАБОТЫ ОБУЧАЮЩИХСЯ ПО ДИСЦИПЛИНЕ

6.1 Общие положения

Знания и навыки, полученные в результате лекций и семинарских занятий, закрепляются и развиваются в результате повторения материала, усвоенного в аудитории, путем чтения текстов и исследовательской литературы (из списков основной и дополнительной литературы) и их анализа.

Самостоятельная работа является важнейшей частью процесса высшего образования. Ее следует осознанно организовать, выделив для этого необходимое время и соответствующим образом организовав рабочее пространство. Важнейшим элементом самостоятельной работы является проработка материалов прошедших занятий (анализ конспектов, чтение рекомендованной литературы) и подготовка к следующим лекциям/семинарским занятиям. Литературу, рекомендованную в программе курса, следует, по возможности, читать в течение всего семестра, концентрируясь на обусловленных программой курса темах.

Существенную часть самостоятельной работы магистранта представляет самостоятельное изучение вспомогательных учебно-методических изданий, лекционных конспектов, интернет-ресурсов и пр. Подготовка к семинарским занятиям является важной формой работы магистранта. Самостоятельная работа может вестись как индивидуально, так и при содействии преподавателя.

6.2 Рекомендации по распределению учебного времени по видам самостоятельной работы и разделам дисциплины

Тема. 1 Введение в R и RStudio

1. Изучение вопросов, представленных в списке тем лекций. Повторение изученного на предыдущих лекциях материала при подготовке к последующим лекциям – 4 часа.

2. Подготовка к занятиям по предложенным для обсуждения вопросам, самостоятельное изучение рекомендованной литературы, повторение материала лекций – 4 часа. Итого: 8 часов.

Тема 2. Основные структуры данных

1. Изучение вопросов, представленных в списке тем лекций. Повторение изученного на предыдущих лекциях материала при подготовке к последующим лекциям – 2 часа.

2. Подготовка к занятиям по предложенным для обсуждения вопросам, самостоятельное изучение рекомендованной литературы, повторение материала лекций – 4 часа. Итого: 6 часов.

Тема 3. Работа с фреймами данных и импортом

1. Изучение вопросов, представленных в списке тем лекций. Повторение изученного на предыдущих лекциях материала при подготовке к последующим лекциям – 2 часа.

2. Подготовка к занятиям по предложенным для обсуждения вопросам, самостоятельное изучение рекомендованной литературы, повторение материала лекций – 4 часа. Итого: 6 часов.

Тема 4. Условные конструкции и циклы

1. Изучение вопросов, представленных в списке тем лекций. Повторение изученного на предыдущих лекциях материала при подготовке к последующим лекциям – 2 часа.

2. Подготовка к занятиям по предложенным для обсуждения вопросам, самостоятельное изучение рекомендованной литературы, повторение материала лекций – 4 часа. Итого: 6 часов.

Тема 5. Функции и модульность кода

1. Изучение вопросов, представленных в списке тем лекций. Повторение изученного на предыдущих лекциях материала при подготовке к последующим лекциям – 2 часа.

2. Подготовка к занятиям по предложенным для обсуждения вопросам, самостоятельное изучение рекомендованной литературы, повторение материала лекций – 4 часа. Итого: 6 часов.

Тема 6. Визуализация данных

1. Изучение вопросов, представленных в списке тем лекций. Повторение изученного на предыдущих лекциях материала при подготовке к последующим лекциям – 2 часа.

2. Подготовка к занятиям по предложенным для обсуждения вопросам, самостоятельное изучение рекомендованной литературы, повторение материала лекций – 4 часа. Итого: 6 часов.

Тема 7. Очистка и подготовка данных к анализу

1. Изучение вопросов, представленных в списке тем лекций. Повторение изученного на предыдущих лекциях материала при подготовке к последующим лекциям – 2 часа.

2. Подготовка к занятиям по предложенным для обсуждения вопросам, самостоятельное изучение рекомендованной литературы, повторение материала лекций – 4 часа. Итого: 6 часов.

6.3 Перечень основных вопросов по изучаемым темам для самостоятельной работы обучающихся по дисциплине

1. Введение в R и RStudio

- Как типы данных в R сравниваются с типами в других языках программирования (например, Python или Java)?
- Какие существуют альтернативные среды разработки для R и чем они отличаются от RStudio?

- Какие возможности есть для автоматизации работы с R через скрипты и командную строку?
 - Как настроить и использовать проекты в RStudio для организации больших исследовательских задач?
2. Основные структуры данных
- Как работают и когда применяются сложные структуры данных, например, дата-фреймы из пакета `tibble`?
 - Какие ограничения и особенности имеют факторы при работе с текстовыми данными?
 - Как объединять и преобразовывать разные типы структур данных (например, преобразовать список в матрицу)?
 - Какие есть альтернативы базовым структурам данных R в других пакетах (например, `data.table`, `dplyr`)?
3. Работа с фреймами данных и импортом
- Какие форматы данных (JSON, XML) можно импортировать в R и с помощью каких пакетов?
 - Какие проблемы возникают при работе с большими данными и как их решить при импорте в R?
 - Как оптимизировать загрузку и хранение данных с учётом производительности?
 - Как использовать функции для пакетной обработки нескольких файлов с данными?
4. Условные конструкции и циклы
- Какие существуют более эффективные альтернативы циклам в R (например, векторизация, функции `apply`)?
 - Как можно комбинировать условные операторы и циклы для сложной логики обработки данных?
 - В каких случаях следует избегать циклов в R, и как это влияет на производительность?
 - Какие существуют встроенные функции для итераций и фильтрации, позволяющие заменить циклы?
5. Функции и модульность кода
- Как реализовать функции с переменным числом аргументов и почему это удобно?
 - Что такое замыкания (closures) в R и как их использовать для создания функций с состоянием?
 - Как документировать функции и использовать пакеты для автоматической генерации документации?
 - Какие практики тестирования функций существуют и как писать тесты в R?
6. Визуализация данных
- Как создавать сложные многокомпонентные графики с `ggplot2` (например, фасетирование, комбинирование нескольких слоев)?
 - Какие возможности анимации графиков доступны в R и с какими пакетами?
 - Как интегрировать визуализации R с веб-приложениями (`shiny`, `flexdashboard`)?
 - Какие есть инструменты для кастомизации и стилизации графиков помимо стандартных функций?
7. Очистка и подготовка данных к анализу
- Как работать с неструктурированными или полуструктурированными данными?
 - Какие методы выявления и устранения выбросов и аномалий существуют?

- Как автоматизировать процесс очистки и обработки данных с помощью пайплайнов и скриптов?
- Какие существуют продвинутое методы трансформации данных, включая работу с временными рядами и текстовыми данными?

6.4 Перечень литературы для самостоятельной работы обучающегося:

1. Дятлов, А. В. Методы многомерного статистического анализа данных в социологии : учебник / А. В. Дятлов, П. Н. Лукичев ; Южный федеральный университет. - Ростов-на-Дону : Издательство Южного федерального университета, 2023. - 236 с. - ISBN 978-5-9275-4265-9. - Текст : электронный. - URL: <https://znanium.ru/catalog/product/2135836>

6.5 Перечень учебно-методического обеспечения для самостоятельной работы

Для обеспечения самостоятельной работы магистрантов по дисциплине «Введение в программирование на R» разработано учебно-методическое обеспечение в составе:

1. Контрольные задания для подготовки к процедурам текущего контроля (п. 7.2 Рабочей программы).
2. Типовые задания для подготовки к промежуточной аттестации (п. 7.4 Рабочей программы).
3. Рекомендуемые основная, дополнительная литература, Интернет-ресурсы и справочные системы (п. 8, 9 Рабочей программы).
4. Рабочая программа дисциплины размещена в электронной информационно-образовательной среде Университета на электронном учебно-методическом ресурсе АНООВО «ЕУСПб» — образовательном портале LMS Sakai — Sakai@EU.

7. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ДЛЯ ПРОВЕДЕНИЯ ТЕКУЩЕЙ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ОБУЧАЮЩИХСЯ ПО ДИСЦИПЛИНЕ

7.1 Показатели, критерии и оценивание компетенций и индикаторов их достижения в процессе текущей аттестации

Информация о содержании и процедуре текущего контроля успеваемости, методике оценивания знаний, умений и навыков обучающегося в ходе текущего контроля доводятся научно-педагогическими работниками Университета до сведения обучающегося на первом занятии по данной дисциплине.

Текущий контроль предусматривает подготовку магистрантов к каждому лабораторному занятию, выполнение домашних заданий, активное слушание на лекциях. Магистрант должен присутствовать на семинарских занятиях, отвечать на поставленные вопросы, показывая, что прочитал разбираемую литературу, представлять содержательные реплики по обсуждаемым вопросам.

Текущий контроль проводится в форме оценивания выполненных домашних заданий, демонстрирующих степень знакомства с дополнительной литературой.

Таблица 5

Показатели, критерии и оценивание компетенций и индикаторов их достижения в процессе текущей аттестации

Наименование тем (разделов)	Коды компетенций	Индикаторы компетенций	Коды ЗУВ (в соот. с Таблицей 1)	Формы текущего контроля успеваемости	Результаты текущего контроля
Введение в R и RStudio	ОПК -1 (ПИ) ОПК -4 (ПИ)	ИД.ОПК-1.1. ИД.ОПК-1.2. ИД.ОПК-4.1. ИД.ОПК-4.2. ИД.ОПК-4.3. ИД.ОПК-4.4.	З (ОПК-1) У (ОПК-1) В (ОПК-1) З (ОПК-4) У (ОПК-4) В (ОПК-4)	Домашнее задание	зачтено/ не зачтено

Наименование тем (разделов)	Коды компетенций	Индикаторы компетенций	Коды ЗУВ (в соот. с Таблицей 1)	Формы текущего контроля успеваемости	Результаты текущего контроля
Основные структуры данных	ОПК -1 (ПИ) ОПК -4 (ПИ)	ИД.ОПК-1.1. ИД.ОПК-1.2. ИД.ОПК-4.1. ИД.ОПК-4.2. ИД.ОПК-4.3. ИД.ОПК-4.4.	З (ОПК-1) У (ОПК-1) В (ОПК-1) З (ОПК-4) У (ОПК-4) В (ОПК-4)	Домашнее задание	зачтено/ не зачтено
Работа с фреймами данных и импортом	ОПК -1 (ПИ) ОПК -4 (ПИ)	ИД.ОПК-1.1. ИД.ОПК-1.2. ИД.ОПК-4.1. ИД.ОПК-4.2. ИД.ОПК-4.3. ИД.ОПК-4.4.	З (ОПК-1) У (ОПК-1) В (ОПК-1) З (ОПК-4) У (ОПК-4) В (ОПК-4)	Домашнее задание	зачтено/ не зачтено
Условные конструкции и циклы	ОПК -1 (ПИ) ОПК -4 (ПИ)	ИД.ОПК-1.1. ИД.ОПК-1.2. ИД.ОПК-4.1. ИД.ОПК-4.2. ИД.ОПК-4.3. ИД.ОПК-4.4.	З (ОПК-1) У (ОПК-1) В (ОПК-1) З (ОПК-4) У (ОПК-4) В (ОПК-4)	Домашнее задание	зачтено/ не зачтено
Функции и модульность кода	ОПК -1 (ПИ) ОПК -4 (ПИ)	ИД.ОПК-1.1. ИД.ОПК-1.2. ИД.ОПК-4.1. ИД.ОПК-4.2. ИД.ОПК-4.3. ИД.ОПК-4.4.	З (ОПК-1) У (ОПК-1) В (ОПК-1) З (ОПК-4) У (ОПК-4) В (ОПК-4)	Домашнее задание	зачтено/ не зачтено
Визуализация данных	ОПК -1 (ПИ) ОПК -4 (ПИ)	ИД.ОПК-1.1. ИД.ОПК-1.2. ИД.ОПК-4.1. ИД.ОПК-4.2. ИД.ОПК-4.3. ИД.ОПК-4.4.	З (ОПК-1) У (ОПК-1) В (ОПК-1) З (ОПК-4) У (ОПК-4) В (ОПК-4)	Домашнее задание	зачтено/ не зачтено
Очистка и подготовка данных к анализу	ОПК -1 (ПИ) ОПК -4 (ПИ)	ИД.ОПК-1.1. ИД.ОПК-1.2. ИД.ОПК-4.1. ИД.ОПК-4.2. ИД.ОПК-4.3. ИД.ОПК-4.4.	З (ОПК-1) У (ОПК-1) В (ОПК-1) З (ОПК-4) У (ОПК-4) В (ОПК-4)	Домашнее задание	зачтено/ не зачтено

Таблица 6

Критерии оценивания

Формы текущего контроля успеваемости	Критерии оценивания
Домашнее задание	Магистрант выполняет работу частично или с существенными недочетами (некорректно сформулирован исследовательский вопрос, не определены основные социальные агенты, некорректно выбраны методы исследования, требования к содержанию, структуре, логике, аргументации, оформлению не выполнены) – не зачтено, Полное и правильное выполнение заданий работы в соответствии с требованиями к содержанию, структуре, логике, аргументации, оформлению с возможным небольшим количеством погрешностей (например, плохо выдержанная структура текста, недостаточная аргументация отдельных тезисов) – зачтено

7.2 Контрольные задания для текущей аттестации

Примерные задания для домашних заданий

Тема 1. Введение в R и RStudio

1. Создайте скрипт, в котором объявите переменные четырех типов: числовой, строковый, логический и комплексный.
2. Выполните для каждой переменной несколько базовых операций: сложение, сравнение, конкатенация.
3. Создайте простой числовой вектор и выполните с ним поэлементные арифметические и логические операции.
4. Для всех объектов определите их класс, тип и длину с помощью встроенных функций.
5. Добавьте комментарии, поясняющие назначение каждой операции.
6. В завершение, создайте блок, в котором производится простейшая агрегация данных: вычисляются сумма, среднее и проверка условий (больше/меньше заданного значения).

Тема 2. Основные структуры данных

1. Создайте: числовой вектор из 15 элементов; список, включающий вектор, строку и логическое значение; матрицу размером 4 на 4, заполненную числами; фактор, построенный на основе категориального вектора с повторяющимися значениями.
2. Проведите индексацию: выберите отдельные элементы из каждого объекта, замените значения, объедините два вектора и создайте новый.
3. Сравните поведение различных структур данных при применении к ним одинаковых операций (сложение, объединение, извлечение).
4. Используйте функции `str`, `summary` и `dim` для анализа структуры объектов. Сопроводите их результат краткими текстовыми выводами.
5. Сравните, в чём практическое различие между списком и фреймом данных.

Тема 3. Работа с фреймами данных и импортом

1. Используйте встроенный набор данных `mtcars`. Просмотрите структуру и содержимое набора: отобразите первые и последние строки, определите число строк и столбцов, выведите названия переменных.
2. Получите доступ к отдельному столбцу и к нескольким выбранным строкам.
3. Выполните фильтрацию строк по двум условиям, связанным с числовыми переменными.
4. Создайте новый столбец, результат которого вычисляется на основе двух других переменных.
5. Отсортируйте данные по одному из столбцов в порядке убывания.
6. Сохраните изменённый датафрейм в отдельный объект.
7. Напишите комментарии, поясняющие каждый шаг обработки.

Тема 4. Условные конструкции и циклы

1. Напишите функцию, которая принимает одно число и возвращает текстовое сообщение, является ли оно положительным, отрицательным или нулём.
2. Сгенерируйте числовой вектор длиной 30 со случайными целыми значениями в диапазоне от -10 до 10.
3. Используя цикл или функцию из семейства `apply`, примените созданную функцию ко всем элементам вектора.
4. Создайте ещё одну функцию, которая принимает числовой вектор и возвращает количество положительных, отрицательных и нулевых значений в виде списка.
5. Напишите обёртку, объединяющую вызов обеих функций и формирующую краткий текстовый отчёт по вектору.
6. Добавьте блок с генерацией нового случайного вектора и вызовом вашей обёртки.

Тема 5. Функции и модульность кода

1. Создайте пользовательскую функцию, которая принимает числовой вектор и возвращает список со следующей информацией: минимальное и максимальное значение, среднее, медиана, стандартное отклонение.
2. Добавьте в функцию проверку: если аргумент не является числовым вектором, выводится сообщение об ошибке.
3. Напишите дополнительную функцию, которая визуализирует полученные сводные статистики в текстовой форме (например, таблица или подписанный вывод).
4. Проверьте работу функций на нескольких различных наборах данных (созданных вручную или встроенных).
5. Объедините всё в один скрипт с логичной структурой и подробными комментариями.

Тема 6. Визуализация данных

1. Используйте встроенный набор данных `mtcars`. Постройте диаграмму рассеяния зависимости расхода топлива от мощности двигателя. Настройте цвет точек по числу цилиндров.
2. Постройте гистограмму распределения расхода топлива, указав количество интервалов и подписи.
3. Постройте `boxplot` сравнения расхода топлива между группами по числу цилиндров.
4. Добавьте заголовки, оси, подписи, измените стиль графиков с использованием функций форматирования.
5. Используя пакет `plotly`, создайте интерактивный график рассеяния на основе того же набора данных, с возможностью отображения значений при наведении.
6. Соберите все графики в один скрипт и прокомментируйте, в каких ситуациях каждый вид графика наиболее информативен.

Тема 7. Очистка и подготовка данных к анализу

1. Используйте встроенный набор данных `airquality`. Просмотрите структуру данных и выявите переменные с пропущенными значениями.
2. Создайте копию набора данных и обработайте пропуски: удалите строки с отсутствующими значениями; в отдельной версии — замените пропуски средними значениями по соответствующим столбцам.
3. Выполните фильтрацию: выберите дни с температурой выше заданного порога и уровнем озона в пределах определённого диапазона.
4. Проведите группировку данных по месяцу и рассчитайте среднее значение температуры и уровня озона.
5. Объедините итоговую таблицу с другой таблицей, где каждому месяцу сопоставлено текстовое название.
6. Сопроводите каждый шаг комментариями и краткими выводами, отражающими смысл проделанных операций.

7.3 Показатели, критерии и оценивание компетенций и индикаторов их достижения в процессе промежуточной аттестации

Форма промежуточной аттестации – зачет с оценкой в форме тестирования.

Перед зачетом с оценкой проводится консультация, на которой преподаватель отвечает на вопросы магистрантов.

В результате промежуточного контроля знаний студенты получают оценку по дисциплине.

Таблица 7

Показатели, критерии и оценивание компетенций и индикаторов их достижения в процессе промежуточной аттестации

Форма промежуточной аттестации/вид промежуточной аттестации	Коды компетенций	Индикаторы компетенций (в соот. с Таблицей 1)	Коды ЗУВ (в соот. с Таблицей 1)	Критерии оценивания	Оценка
зачет с оценкой / тестирование	ОПК-1 (ПИ) ОПК-4 (ПИ)	ИД.ОПК-1.1. ИД.ОПК-1.2. ИД.ОПК-4.1. ИД.ОПК-4.2. ИД.ОПК-4.3. ИД.ОПК-4.4.	З (ОПК-1) У (ОПК-1) В (ОПК-1) З (ОПК-4) У (ОПК-4) В (ОПК-4)	100-81% правильных ответов	Зачтено, отлично
				80-61% правильных ответов	Зачтено, хорошо
				60-41% правильных ответов	Зачтено, удовлетворительно
				40-0% правильных ответов	Не зачтено, неудовлетворительно

Результаты сдачи промежуточной аттестации по направлениям подготовки уровня магистратуры оцениваются в соответствии с Положением о формах, периодичности и порядке организации и проведения текущего контроля успеваемости и промежуточной аттестации обучающихся в АНООВО «ЕУСПб» следующим образом согласно таблице 7а.

Таблица 7а

Система оценки знаний обучающихся

Пятибалльная (стандартная) система	Стобалльная система оценки	Бинарная система оценки
5 (отлично)	100-81	зачтено
4 (хорошо)	80-61	
3 (удовлетворительно)	60-41	
2 (неудовлетворительно)	40 и менее	не зачтено

Результаты промежуточного контроля по дисциплине, выраженные в оценках «зачтено, удовлетворительно», «зачтено, хорошо», «зачтено, отлично» показывают уровень сформированности у обучающегося компетенций по дисциплине в соответствии с картами компетенций образовательной программы «Вычислительная социология» по направлению подготовки 39.04.01 Социология (уровень магистратуры).

Результаты промежуточного контроля по дисциплине, выраженные в оценке «не зачтено, неудовлетворительно», показывают не сформированность у обучающегося компетенций по дисциплине в соответствии с картами компетенций образовательной программы «Вычислительная социология» по направлению подготовки 39.04.01 Социология (уровень магистратуры).

7.4 Типовые задания к промежуточной аттестации

Требования к тестам

Тест включает 25 вопросов по всем компетенциям дисциплины, 10 из них вопросы закрытого типа, 5 – комбинированного типа, 10 – открытого типа, все вопросы разного уровня сложности.

Тест оценивается в баллах в соответствии со следующими критериями:

Задания закрытого типа

Базовый уровень сложности: задание считается выполненным верно, если ответ полностью совпадает с эталоном ответа: каждый символ в ответе стоит на своём месте - 1 балл; ответ отличен от эталонного - 0 баллов.

Повышенный уровень сложности: задание считается выполненным верно, если ответ полностью совпадает с эталоном ответа: каждый символ в ответе стоит на своём месте, лишние символы в ответе отсутствуют - 2 балл; если на любой одной позиции ответа

записан не тот символ, который представлен в эталоне ответа - 1 балл; во всех других случаях выставляется 0 баллов

Комбинированные задания

Базовый уровень сложности: задание считается выполненным верно, если ответ полностью совпадает с эталоном ответа: каждый символ в ответе стоит на своём месте, обоснование по смыслу соответствует эталонному (допускаются различные формулировки ответа, не искажающие его смысла) - 1 балл; ответ отличен от эталонного - 0 баллов.

Повышенный уровень сложности: задание считается выполненным верно, если ответ полностью совпадает с эталоном ответа: каждый символ в ответе стоит на своём месте, обоснование по смыслу соответствует эталонному (допускаются различные формулировки ответа, не искажающие его смысла) - 2 балла; дан верный ответ, обоснование отсутствует или приведено неверно – 1 балл; во всех остальных случаях - 0 баллов.

Задания открытого типа

Повышенный уровень сложности: ответ соответствует эталонному (допускаются различные формулировки ответа, не искажающие его смысла); правильно названы все запрашиваемые составляющие вопроса, даны верные обоснования - 2 балла; ответ имеет незначительные отклонения от эталонного, правильно названы на все запрашиваемые составляющие вопроса, но для названных даны верные обоснования - 1 балл; ответ значительно отличается от эталонного, имеются фактические ошибки, искажающие его смысл или ответ сформулирован неверно или не сформулирован - 0 баллов.

Высокий уровень сложности: магистрант демонстрирует умение применять знания в нестандартной ситуации, решать нетиповые задачи, приводит корректные обоснования и доказательства, ответ полный, в ответе отсутствуют фактические ошибки, изложение связное, структура прозрачная, логика изложения прослеживается - 3 балла; ответ значительно отличается от эталонного, имеются фактические ошибки, искажающие его смысл или ответ сформулирован неверно или не сформулирован - 0 баллов.

Итоговый балл за тест рассчитывается по формуле:

$$F = \frac{100}{K} * \left(\frac{x_1}{k_1} + \frac{x_2}{k_2} + \dots + \frac{x_n}{k_n} \right),$$

где F – итоговое количество баллов за тест,

K – количество осваиваемых в рамках дисциплины компетенций,

k_n – максимально возможное количество баллов за вопросы по компетенции,

x_n – количество баллов, набранное магистрантом, за правильные ответы на вопросы по соответствующей компетенции.

Примеры тестовых заданий для промежуточной аттестации

ОПК-1 (ПИ) Способен самостоятельно приобретать, развивать и применять математические, естественнонаучные, социально-экономические и профессиональные знания для решения нестандартных задач, в том числе в новой или незнакомой среде и в междисциплинарном контексте

Задания закрытого типа (базовый уровень сложности)

Вопрос 1

Инструкция: Выберите один правильный ответ.

Вопрос: Какой символ в R используется для присваивания значения переменной?

а) =

б) ==

в) <-

г) :=

Поле для ответа: _____

Задания закрытого типа (повышенный уровень сложности)

Вопрос 1

Инструкция: Установите соответствие между типом объекта и примером в R.

Тип объекта	Пример
а) Вектор	3. c(1, 2, 3, 4)
б) Матрица	1. matrix(1:6, 2, 3)
в) Data frame	4. data.frame(x=1:3, y=c("a","b","c"))
г) Список	2. list(a=1, b="x")

Поле для ответа: _____

Задания комбинированного типа (повышенный уровень сложности)

Вопрос 1

Инструкция: Выберите единственный правильный ответ и приведите обоснование

Вопрос: Вам необходимо передать табличные данные в другой пакет для визуальной выдачи и построения графиков. Какой из операторов будет наиболее приемлимым и почему?

- а) =
- б) <-
- в) ==
- г) %>%

Поле для ответа: _____

Обоснование ответа: _____

Вопрос 2

Инструкция: Выберите единственный правильный ответ и приведите обоснование.

Вопрос: Какой результат работы функции seq(1, 5, by=2)?

- а) 1 2 3 4 5
- б) 1 3 5
- в) 1 4 5
- г) 2 4 6

Поле для ответа: _____

Обоснование: _____

Задания открытого типа (высокий уровень сложности)

Вопрос 1

Инструкция: Ответьте письменно, приведите рассуждение.

Задание: В чем преимущества использования встроенных наборов данных в R для обучения и отладки кода?

Поле для ответа: _____

ОПК-4 (ПИ) Способен применять на практике новые научные принципы и методы исследований

Задания закрытого типа (базовый уровень сложности)

Вопрос 1

Инструкция: Выберите один правильный ответ.

Вопрос: Что делает функция ggplot(data, aes(x, y)) + geom_line()?

- а) Создает точечную диаграмму
- б) Создает линейный график

- в) Создает гистограмму
 г) Создает круговую диаграмму
 Поле для ответа: _____

Задания закрытого типа (повышенный уровень сложности)

Вопрос 1

Инструкция: Установите соответствие между типом индексации и примером.

Тип индексации	Пример
а) Векторная	4. x[2:4]
б) Логическая	2. x[x > 0]
в) Именованная	1. df[,"var1"]
г) Комбинированная	3. df[df\$var2 > 5, "var1"]

Поле для ответа: _____

Задания комбинированного типа (повышенный уровень сложности)

Вопрос 1

Инструкция: Выберите единственный правильный ответ и приведите обоснование ответа

Вопрос: Какая функция возвращает структуру объекта, показывая типы и размеры?

- а) str()
 б) class()
 в) summary()
 г) typeof()

Поле для ответа: _____

Обоснование: _____

Вопрос 2

Инструкция: Выберите единственный правильный ответ и приведите обоснование в сравнении с другими функциями и их функционалом.

Вопрос: Что делает функция filter() из пакета dplyr?

- а) Отбирает строки по условию
 б) Отбирает столбцы по имени
 в) Объединяет две таблицы
 г) Сортирует таблицу

Поле для ответа: _____

Обоснование: _____

Задания открытого типа (высокий уровень сложности)

Инструкция: Дайте развернутый ответ на поставленный вопрос.

Задание: Почему важно настраивать подписи и легенды на графиках?

Поле для ответа: _____

7.5 Средства оценки индикаторов достижения компетенций

Таблица 8

Средства оценки индикаторов достижения компетенций

Коды компетенций	Индикаторы компетенций (в соот.с Таблицей 1)	Средства оценки (в соот. с Таблицами 5, 7)
ОПК-1 (ПИ)	ИД.ОПК-1.1.	Домашнее задание, тест

Коды компетенций	Индикаторы компетенций (в соот.с Таблицей 1)	Средства оценки (в соот. с Таблицами 5, 7)
	ИД.ОПК-1.2.	
ОПК-4 (ПИ)	ИД.ОПК-4.1. ИД.ОПК-4.2. ИД.ОПК-4.3. ИД.ОПК-4.4.	Домашнее задание, тест

Таблица 9

Описание средств оценки индикаторов достижения компетенций

Средства оценки (в соот. с Таблицами 5, 7)	Рекомендованный план выполнения работы
Домашнее задание	<p>Магистрант в ходе подготовки и выполнения домашнего задания показывает наличие практической базы знаний в рамках дисциплины, необходимой для выполнения следующих действий в области профессиональной деятельности:</p> <p>Делать обоснованный выбор методов разработки алгоритмов и программных средств для решения профессиональных задач, в том числе с использованием современных информационных технологий</p> <p>На основе современных теорий и концепций обосновывать актуальность, анализировать, применять и разрабатывать предложения и рекомендации по использованию новых научных принципов и методов исследований в профессиональной области знаний</p>
Тест	<p>Магистрант в ходе подготовки и выполнения тестирования показывает наличие практической базы знаний в рамках дисциплины, необходимой для выполнения следующих действий в области профессиональной деятельности:</p> <p>Делать обоснованный выбор методов разработки алгоритмов и программных средств для решения профессиональных задач, в том числе с использованием современных информационных технологий</p> <p>На основе современных теорий и концепций обосновывать актуальность, анализировать, применять и разрабатывать предложения и рекомендации по использованию новых научных принципов и методов исследований в профессиональной области знаний</p>

8. ОСНОВНАЯ И ДОПОЛНИТЕЛЬНАЯ ЛИТЕРАТУРА

8.1. Основная литература

1. **Золотарюк, А. В.** Язык и среда программирования R : учебное пособие / А.В. Золотарюк. — Москва : ИНФРА-М, 2023. — 162 с. — (Высшее образование). — DOI 10.12737/textbook_5b8fdb0bd795c4.69435980. - ISBN 978-5-16-018723-5. - Текст : электронный. - URL: <https://znanium.com/catalog/product/2049696>

2. **Баяк, О. А.** Практикум по анализу данных на языках Python и R : учебное пособие / О. А. Баяк, М. Р. Исаева, М. О. Самсонкин. - Москва : Прометей, 2023. - 100 с. - ISBN 978-5-00172-356-1. - Текст : электронный. - URL: <https://znanium.ru/catalog/product/2124862>

8.2 Дополнительная литература

1. Дятлов, А. В. Методы многомерного статистического анализа данных в социологии : учебник / А. В. Дятлов, П. Н. Лукичев ; Южный федеральный университет. - Ростов-на-Дону : Издательство Южного федерального университета, 2023. - 236 с. - ISBN 978-5-9275-4265-9. - Текст : электронный. - URL: <https://znanium.ru/catalog/product/2135836>

9. ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ, ИСПОЛЬЗУЕМЫЕ ПРИ ОСУЩЕСТВЛЕНИИ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА

9.1 Программное обеспечение

1. OS Microsoft Windows (OVS OS Platform)
2. MS Office (OVS Office Platform)
3. Adobe Acrobat Professional 11.0 MLP AOO License RU
4. Adobe CS5.5 Design Standart Win IE EDU CLP
5. ABBYY FineReader 11 Corporate Edition
6. ABBYY Lingvo x5
7. Adobe Acrobat Reader DC /Pro – бесплатно
8. Google Chrome – бесплатно
9. Opera – бесплатно
10. Mozilla – бесплатно
11. VLC – бесплатно
12. R – бесплатно
13. Яндекс Браузер – бесплатно
14. Интегрированная среда разработки RStudio (бесплатно)

9.2 Перечень информационно-справочных систем и профессиональных баз данных информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины:

Информационно-справочные системы

1. Гарант.Ру. Информационно-правовой портал: <http://www.garant.ru>
2. Информационная система «Единое окно доступа к образовательным ресурсам»: <http://window.edu.ru/>
3. Открытое образование. Ассоциация «Национальная платформа открытого образования»: <http://npoed.ru>
4. Официальная Россия. Сервер органов государственной власти Российской Федерации: <http://www.gov.ru>
5. Официальный интернет-портал правовой информации. Государственная система правовой информации: <http://pravo.gov.ru>
6. Правовой сайт КонсультантПлюс: <http://www.consultant.ru/sys>
7. Российское образование. Федеральный портал: <http://www.edu.ru>

Профессиональные базы данных информационно-телекоммуникационной сети «Интернет»:

1. Национальная электронная библиотека НЭБ: <http://www.rusneb.ru>
2. Президентская библиотека: <http://www.prlib.ru>
3. Российская государственная библиотека: <http://www.rsl.ru/>
4. Российская национальная библиотека: <http://www.nlr.ru/poisk/>

9.3 Лицензионные электронные ресурсы библиотеки Университета

Профессиональные базы данных:

Полный перечень доступных обучающимся профессиональных баз данных представлен на официальном сайте Университета <https://eusp.org/library/electronic-resources>, включая следующие базы данных:

1. **eLIBRARY.RU** — Российский информационно-аналитический портал в области науки, технологии, медицины и образования, содержащий рефераты и полные тексты научных статей и публикаций, наукометрическая база данных: <http://elibrary.ru>;
2. **Университетская информационная система РОССИЯ** — база электронных ресурсов для учебных программ и исследовательских проектов в области социально-гуманитарных наук: <http://www.uisrussia.msu.ru/>;

3. Электронные журналы по подписке (текущие номера научных зарубежных журналов).

Электронные библиотечные системы:

1. **Znanium.com** – Электронная библиотечная система (ЭБС) – <http://znanium.com/>;

2. Университетская библиотека онлайн – Электронная библиотечная система (ЭБС) – <http://biblioclub.ru/>

9.4 Электронная информационно-образовательная среда Университета

Образовательный процесс по дисциплине поддерживается средствами электронной информационно-образовательной среды Университета, которая включает в себя электронный учебно-методический ресурс АНООВО «ЕУСПб» — образовательный портал LMS Sakai — Sakai@EU, лицензионные электронные ресурсы библиотеки Университета, официальный сайт Университета (Европейский университет в Санкт-Петербурге [<https://euspr.org/>]), локальную сеть и корпоративную электронную почту Университета, и обеспечивает:

- доступ к учебным планам, рабочим программам дисциплин (модулей), практик и к изданиям электронных библиотечных систем и электронным образовательным ресурсам, указанным в рабочих программах;
- фиксацию хода образовательного процесса, результатов промежуточной аттестации и результатов освоения основной образовательной программы;
- формирование электронного портфолио обучающегося, в том числе сохранение работ обучающегося, рецензий и оценок за эти работы со стороны любых участников образовательного процесса;
- взаимодействие между участниками образовательного процесса, в том числе синхронное и (или) асинхронное взаимодействие посредством сети «Интернет» (электронной почты и т.д.).

Каждый обучающийся в течение всего периода обучения обеспечен индивидуальным неограниченным доступом к электронным ресурсам библиотеки Университета, содержащей издания учебной, учебно-методической и иной литературы по изучаемой дисциплине

10. МАТЕРИАЛЬНО-ТЕХНИЧЕСКАЯ БАЗА, НЕОБХОДИМАЯ ДЛЯ ОСУЩЕСТВЛЕНИЯ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА

В ходе реализации образовательного процесса используются специализированные многофункциональные аудитории для проведения занятий лекционного типа, занятий семинарского типа (практических занятий, лабораторных работ), групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, укомплектованные специализированной мебелью и техническими средствами обучения, служащими для представления учебной информации большой аудитории.

Проведение занятий лекционного типа обеспечивается демонстрационным оборудованием.

Помещения для самостоятельной работы оснащены компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду организации.

Для лиц с ограниченными возможностями здоровья и инвалидов предоставляется возможность присутствия в аудитории вместе с ними ассистента (помощника). Для слабовидящих предоставляется возможность увеличения текста на экране ПК. Для самостоятельной работы лиц с ограниченными возможностями здоровья в помещении для самостоятельной работы организовано одно место (ПК) с возможностями бесконтактного ввода информации и управления компьютером (специализированное лицензионное

программное обеспечение – Camera Mouse, веб камера). Библиотека университета предоставляет удаленный доступ к электронным ресурсам библиотеки Университета с возможностями для слабовидящих увеличения текста на экране ПК. Лица с ограниченными возможностями здоровья могут при необходимости воспользоваться имеющимся в университете креслом-коляской. В учебном корпусе имеется адаптированный лифт. На первом этаже оборудован специализированный туалет. У входа в здание университета для инвалидов оборудована специальная кнопка, входная среда обеспечена информационной доской о режиме работы университета, выполненной рельефно-точечным тактильным шрифтом (азбука Брайля).

ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ПО ДИСЦИПЛИНЕ
«Введение в программирование на R»

**ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ДЛЯ ПРОВЕДЕНИЯ ТЕКУЩЕЙ И
ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ОБУЧАЮЩИХСЯ ПО ДИСЦИПЛИНЕ**

1 Показатели, критерии и оценивание компетенций и индикаторов их достижения в процессе текущей аттестации

Информация о содержании и процедуре текущего контроля успеваемости, методике оценивания знаний, умений и навыков обучающегося в ходе текущего контроля доводятся научно-педагогическими работниками Университета до сведения обучающегося на первом занятии по данной дисциплине.

Текущий контроль предусматривает подготовку магистрантов к каждому лабораторному занятию, выполнение домашних заданий, активное слушание на лекциях. Магистрант должен присутствовать на семинарских занятиях, отвечать на поставленные вопросы, показывая, что прочитал разбираемую литературу, представлять содержательные реплики по обсуждаемым вопросам.

Текущий контроль проводится в форме оценивания выполненных домашних заданий, демонстрирующих степень знакомства с дополнительной литературой.

Таблица 1

**Показатели, критерии и оценивание компетенций и индикаторов их
достижения в процессе текущей аттестации**

Наименование тем (разделов)	Коды компетенций	Индикаторы компетенций	Коды ЗУВ (в соот. с Таблицей 1)	Формы текущего контроля успеваемости	Результаты текущего контроля
Введение в R и RStudio	ОПК -1 (ПИ) ОПК -4 (ПИ)	ИД.ОПК-1.1. ИД.ОПК-1.2. ИД.ОПК-4.1. ИД.ОПК-4.2. ИД.ОПК-4.3. ИД.ОПК-4.4.	З (ОПК-1) У (ОПК-1) В (ОПК-1) З (ОПК-4) У (ОПК-4) В (ОПК-4)	Домашнее задание	зачтено/ не зачтено
Основные структуры данных	ОПК -1 (ПИ) ОПК -4 (ПИ)	ИД.ОПК-1.1. ИД.ОПК-1.2. ИД.ОПК-4.1. ИД.ОПК-4.2. ИД.ОПК-4.3. ИД.ОПК-4.4.	З (ОПК-1) У (ОПК-1) В (ОПК-1) З (ОПК-4) У (ОПК-4) В (ОПК-4)	Домашнее задание	зачтено/ не зачтено
Работа с фреймами данных и импортом	ОПК -1 (ПИ) ОПК -4 (ПИ)	ИД.ОПК-1.1. ИД.ОПК-1.2. ИД.ОПК-4.1. ИД.ОПК-4.2. ИД.ОПК-4.3. ИД.ОПК-4.4.	З (ОПК-1) У (ОПК-1) В (ОПК-1) З (ОПК-4) У (ОПК-4) В (ОПК-4)	Домашнее задание	зачтено/ не зачтено
Условные конструкции и циклы	ОПК -1 (ПИ) ОПК -4 (ПИ)	ИД.ОПК-1.1. ИД.ОПК-1.2. ИД.ОПК-4.1. ИД.ОПК-4.2. ИД.ОПК-4.3. ИД.ОПК-4.4.	З (ОПК-1) У (ОПК-1) В (ОПК-1) З (ОПК-4) У (ОПК-4) В (ОПК-4)	Домашнее задание	зачтено/ не зачтено
Функции и модульность кода	ОПК -1 (ПИ) ОПК -4 (ПИ)	ИД.ОПК-1.1. ИД.ОПК-1.2. ИД.ОПК-4.1. ИД.ОПК-4.2. ИД.ОПК-4.3. ИД.ОПК-4.4.	З (ОПК-1) У (ОПК-1) В (ОПК-1) З (ОПК-4) У (ОПК-4) В (ОПК-4)	Домашнее задание	зачтено/ не зачтено
Визуализация данных	ОПК -1 (ПИ) ОПК -4 (ПИ)	ИД.ОПК-1.1. ИД.ОПК-1.2. ИД.ОПК-4.1. ИД.ОПК-4.2. ИД.ОПК-4.3. ИД.ОПК-4.4.	З (ОПК-1) У (ОПК-1) В (ОПК-1) З (ОПК-4) У (ОПК-4) В (ОПК-4)	Домашнее задание	зачтено/ не зачтено

Наименование тем (разделов)	Коды компетенций	Индикаторы компетенций	Коды ЗУВ (в соот. с Таблицей 1)	Формы текущего контроля успеваемости	Результаты текущего контроля
Очистка и подготовка данных к анализу	ОПК -1 (ПИ) ОПК -4 (ПИ)	ИД.ОПК-1.1. ИД.ОПК-1.2. ИД.ОПК-4.1. ИД.ОПК-4.2. ИД.ОПК-4.3. ИД.ОПК-4.4.	З (ОПК-1) У (ОПК-1) В (ОПК-1) З (ОПК-4) У (ОПК-4) В (ОПК-4)	Домашнее задание	зачтено/ не зачтено

Таблица 2

Критерии оценивания

Формы текущего контроля успеваемости	Критерии оценивания
Домашнее задание	Магистрант выполняет работу частично или с существенными недочетами (некорректно сформулирован исследовательский вопрос, не определены основные социальные агенты, некорректно выбраны методы исследования, требования к содержанию, структуре, логике, аргументации, оформлению не выполнены) – не зачтено, Полное и правильное выполнение заданий работы в соответствии с требованиями к содержанию, структуре, логике, аргументации, оформлению с возможным небольшим количеством погрешностей (например, плохо выдержанная структура текста, недостаточная аргументация отдельных тезисов) – зачтено

2 Контрольные задания для текущей аттестации

Задания для домашних заданий

Тема 1. Введение в R и RStudio

1. Создайте скрипт, в котором объявите переменные четырех типов: числовой, строковый, логический и комплексный.
2. Выполните для каждой переменной несколько базовых операций: сложение, сравнение, конкатенация.
3. Создайте простой числовой вектор и выполните с ним поэлементные арифметические и логические операции.
4. Для всех объектов определите их класс, тип и длину с помощью встроенных функций.
5. Добавьте комментарии, поясняющие назначение каждой операции.
6. В завершение, создайте блок, в котором производится простейшая агрегация данных: вычисляются сумма, среднее и проверка условий (больше/меньше заданного значения).

Тема 2. Основные структуры данных

1. Создайте: числовой вектор из 15 элементов; список, включающий вектор, строку и логическое значение; матрицу размером 4 на 4, заполненную числами; фактор, построенный на основе категориального вектора с повторяющимися значениями.
2. Проведите индексацию: выберите отдельные элементы из каждого объекта, замените значения, объедините два вектора и создайте новый.
3. Сравните поведение различных структур данных при применении к ним одинаковых операций (сложение, объединение, извлечение).
4. Используйте функции `str`, `summary` и `dim` для анализа структуры объектов. Сопроводите их результат краткими текстовыми выводами.
5. Сравните, в чём практическое различие между списком и фреймом данных.

Тема 3. Работа с фреймами данных и импортом

1. Используйте встроенный набор данных `mtcars`. Просмотрите структуру и содержимое набора: отобразите первые и последние строки, определите число строк и столбцов, выведите названия переменных.
2. Получите доступ к отдельному столбцу и к нескольким выбранным строкам.
3. Выполните фильтрацию строк по двум условиям, связанным с числовыми переменными.
4. Создайте новый столбец, результат которого вычисляется на основе двух других переменных.
5. Отсортируйте данные по одному из столбцов в порядке убывания.
6. Сохраните изменённый датафрейм в отдельный объект.
7. Напишите комментарии, поясняющие каждый шаг обработки.

Тема 4. Условные конструкции и циклы

1. Напишите функцию, которая принимает одно число и возвращает текстовое сообщение, является ли оно положительным, отрицательным или нулём.
2. Сгенерируйте числовой вектор длиной 30 со случайными целыми значениями в диапазоне от -10 до 10.
3. Используя цикл или функцию из семейства `apply`, примените созданную функцию ко всем элементам вектора.
4. Создайте ещё одну функцию, которая принимает числовой вектор и возвращает количество положительных, отрицательных и нулевых значений в виде списка.
5. Напишите обёртку, объединяющую вызов обеих функций и формирующую краткий текстовый отчёт по вектору.
6. Добавьте блок с генерацией нового случайного вектора и вызовом вашей обёртки.

Тема 5. Функции и модульность кода

1. Создайте пользовательскую функцию, которая принимает числовой вектор и возвращает список со следующей информацией: минимальное и максимальное значение, среднее, медиана, стандартное отклонение.
2. Добавьте в функцию проверку: если аргумент не является числовым вектором, выводится сообщение об ошибке.
3. Напишите дополнительную функцию, которая визуализирует полученные сводные статистики в текстовой форме (например, таблица или подписанный вывод).
4. Проверьте работу функций на нескольких различных наборах данных (созданных вручную или встроенных).
5. Объедините всё в один скрипт с логичной структурой и подробными комментариями.

Тема 6. Визуализация данных

1. Используйте встроенный набор данных `mtcars`. Постройте диаграмму рассеяния зависимости расхода топлива от мощности двигателя. Настройте цвет точек по числу цилиндров.
2. Постройте гистограмму распределения расхода топлива, указав количество интервалов и подписи.
3. Постройте `boxplot` сравнения расхода топлива между группами по числу цилиндров.
4. Добавьте заголовки, оси, подписи, измените стиль графиков с использованием функций форматирования.
5. Используя пакет `plotly`, создайте интерактивный график рассеяния на основе того же набора данных, с возможностью отображения значений при наведении.
6. Соберите все графики в один скрипт и прокомментируйте, в каких ситуациях каждый вид графика наиболее информативен.

Тема 7. Очистка и подготовка данных к анализу

1. Используйте встроенный набор данных `airquality`. Просмотрите структуру данных и выявите переменные с пропущенными значениями.
2. Создайте копию набора данных и обработайте пропуски: удалите строки с отсутствующими значениями; в отдельной версии — замените пропуски средними значениями по соответствующим столбцам.
3. Выполните фильтрацию: выберите дни с температурой выше заданного порога и уровнем озона в пределах определённого диапазона.
4. Проведите группировку данных по месяцу и рассчитайте среднее значение температуры и уровня озона.
5. Объедините итоговую таблицу с другой таблицей, где каждому месяцу сопоставлено текстовое название.
6. Сопроводите каждый шаг комментариями и краткими выводами, отражающими смысл проделанных операций.

3 Показатели, критерии и оценивание компетенций и индикаторов их достижения в процессе промежуточной аттестации

Форма промежуточной аттестации – зачет с оценкой в форме тестирования.

Перед зачетом с оценкой проводится консультация, на которой преподаватель отвечает на вопросы магистрантов.

В результате промежуточного контроля знаний студенты получают оценку по дисциплине.

Таблица 3

Показатели, критерии и оценивание компетенций и индикаторов их достижения в процессе промежуточной аттестации

Форма промежуточной аттестации/вид промежуточной аттестации	Коды компетенций	Индикаторы компетенций (в соот. с Таблицей 1)	Коды ЗУВ (в соот. с Таблицей 1)	Критерии оценивания	Оценка
зачет с оценкой / тестирование	ОПК -1 (ПИ) ОПК -4 (ПИ)	ИД.ОПК-1.1. ИД.ОПК-1.2. ИД.ОПК-4.1. ИД.ОПК-4.2. ИД.ОПК-4.3. ИД.ОПК-4.4.	З (ОПК-1) У (ОПК-1) В (ОПК-1) З (ОПК-4) У (ОПК-4) В (ОПК-4)	100-81% правильных ответов	Зачтено, отлично
				80-61% правильных ответов	Зачтено, хорошо
				60-41% правильных ответов	Зачтено, удовлетворительно
				40-0% правильных ответов	Не зачтено, неудовлетворительно

Результаты сдачи промежуточной аттестации по направлениям подготовки уровня магистратуры оцениваются в соответствии с Положением о формах, периодичности и порядке организации и проведения текущего контроля успеваемости и промежуточной аттестации обучающихся в АНООВО «ЕУСПб» следующим образом согласно таблице 3а.

Таблица 3а

Система оценки знаний обучающихся

Пятибалльная (стандартная) система	Стобалльная система оценки	Бинарная система оценки
5 (отлично)	100-81	зачтено
4 (хорошо)	80-61	
3 (удовлетворительно)	60-41	
2 (неудовлетворительно)	40 и менее	не зачтено

Результаты промежуточного контроля по дисциплине, выраженные в оценках «зачтено, удовлетворительно», «зачтено, хорошо», «зачтено, отлично» показывают уровень сформированности у обучающегося компетенций по дисциплине в соответствии с картами компетенций образовательной программы «Вычислительная социология» по направлению подготовки 39.04.01 Социология (уровень магистратуры).

Результаты промежуточного контроля по дисциплине, выраженные в оценке «не зачтено, неудовлетворительно», показывают не сформированность у обучающегося компетенций по дисциплине в соответствии с картами компетенций образовательной программы «Вычислительная социология» по направлению подготовки 39.04.01 Социология (уровень магистратуры).

4 Задания к промежуточной аттестации

Требования к тестам

Тест включает 25 вопросов по всем компетенциям дисциплины, 10 из них вопросы закрытого типа, 5 – комбинированного типа, 10 – открытого типа, все вопросы разного уровня сложности.

Тест оценивается в баллах в соответствии со следующими критериями:

Задания закрытого типа

Базовый уровень сложности: задание считается выполненным верно, если ответ полностью совпадает с эталоном ответа: каждый символ в ответе стоит на своём месте - 1 балл; ответ отличен от эталонного - 0 баллов.

Повышенный уровень сложности: задание считается выполненным верно, если ответ полностью совпадает с эталоном ответа: каждый символ в ответе стоит на своём месте, лишние символы в ответе отсутствуют - 2 балла; если на любой одной позиции ответа записан не тот символ, который представлен в эталоне ответа - 1 балл; во всех других случаях выставляется 0 баллов

Комбинированные задания

Базовый уровень сложности: задание считается выполненным верно, если ответ полностью совпадает с эталоном ответа: каждый символ в ответе стоит на своём месте, обоснование по смыслу соответствует эталонному (допускаются различные формулировки ответа, не искажающие его смысла) - 1 балл; ответ отличен от эталонного - 0 баллов.

Повышенный уровень сложности: задание считается выполненным верно, если ответ полностью совпадает с эталоном ответа: каждый символ в ответе стоит на своём месте, обоснование по смыслу соответствует эталонному (допускаются различные формулировки ответа, не искажающие его смысла) - 2 балла; дан верный ответ, обоснование отсутствует или приведено неверно – 1 балл; во всех остальных случаях - 0 баллов.

Задания открытого типа

Повышенный уровень сложности: ответ соответствует эталонному (допускаются различные формулировки ответа, не искажающие его смысла); правильно названы все запрашиваемые составляющие вопроса, даны верные обоснования - 2 балла; ответ имеет незначительные отклонения от эталонного, правильно названы на все запрашиваемые составляющие вопроса, но для названных даны верные обоснования - 1 балл; ответ значительно отличается от эталонного, имеются фактические ошибки, искажающие его смысл или ответ сформулирован неверно или не сформулирован - 0 баллов.

Высокий уровень сложности: магистрант демонстрирует умение применять знания в нестандартной ситуации, решать нетиповые задачи, приводит корректные обоснования и доказательства, ответ полный, в ответе отсутствуют фактические ошибки, изложение

связное, структура прозрачная, логика изложения прослеживается - 3 балла; ответ значительно отличается от эталонного, имеются фактические ошибки, искажающие его смысл или ответ сформулирован неверно или не сформулирован - 0 баллов.

Итоговый балл за тест рассчитывается по формуле:

$$F = \frac{100}{K} * \left(\frac{x_1}{k_1} + \frac{x_2}{k_2} + \dots + \frac{x_n}{k_n} \right),$$

где F – итоговое количество баллов за тест,

K – количество осваиваемых в рамках дисциплины компетенций,

k_n – максимально возможное количество баллов за вопросы по компетенции,

x_n – количество баллов, набранное магистрантом, за правильные ответы на вопросы по соответствующей компетенции.

Задания к промежуточной аттестации

ОПК-1 (ПИ) Способен самостоятельно приобретать, развивать и применять математические, естественнонаучные, социально-экономические и профессиональные знания для решения нестандартных задач, в том числе в новой или незнакомой среде и в междисциплинарном контексте

Задания закрытого типа (базовый уровень сложности)

Задание 1

Инструкция: Выберите Правильный ответ из предложенных.

Вопрос: Какой оператор в R используется для присваивания значения переменной?

Варианты ответа:

1. ==
2. =
3. <-
4. ->

Правильный ответ:

Задание 2

Инструкция: Выберите Правильный ответ из предложенных.

Вопрос: Какая функция в R позволяет получить список всех объектов в рабочей среде?

Варианты ответа:

1. ls()
2. dir()
3. getwd()
4. objects()

Правильный ответ:

Задание 3

Инструкция: Выберите Правильный ответ из предложенных.

Вопрос: Какой тип данных в R используется для хранения текстовых значений?

Варианты ответа:

1. numeric
2. character
3. logical
4. integer

Правильный ответ:

Задание 4

Инструкция: Выберите Правильный ответ из предложенных.

Вопрос: Какая функция используется для чтения данных из CSV-файла в R?

Варианты ответа:

1. read.csv()
2. load.csv()
3. import.csv()
4. open.csv()

Правильный ответ:

Задание 5

Инструкция: Выберите Правильный ответ из предложенных.

Вопрос: Какой пакет в R является наиболее популярным для создания статических графиков?

Варианты ответа:

1. plotly
2. ggplot2
3. lattice
4. base

Правильный ответ:

Задание 6

Инструкция: Выберите Правильный ответ из предложенных.

Вопрос: Какая функция позволяет получить структуру объекта в R?

Варианты ответа:

1. summary()
2. str()
3. dim()
4. class()

Правильный ответ:

Задание 7

Инструкция: Выберите Правильный ответ из предложенных.

Вопрос: Какой оператор используется для выполнения логического "И" в R?

Варианты ответа:

1. &
2. &&
3. |
4. ||

Правильный ответ:

Задание 8

Инструкция: Выберите Правильный ответ из предложенных.

Вопрос: Какая функция в R позволяет создать последовательность чисел?

Варианты ответа:

1. rep()
2. seq()
3. c()
4. paste()

Правильный ответ:

Задание 9

Инструкция: Выберите Правильный ответ из предложенных.

Вопрос: Какой символ используется для комментирования кода в R?

Варианты ответа:

1. #
2. //
3. /* */
4. –

Правильный ответ:

Задание 10

Инструкция: Выберите Правильный ответ из предложенных.

Вопрос: Какая функция в R позволяет объединить два вектора в один?

Варианты ответа:

1. merge()
2. cbind()
3. c()
4. append()

Правильный ответ:

Задания закрытого типа (повышенный уровень сложности):

Задание 11

Инструкция: Сопоставьте функцию в R с её описанием.

Вопрос:

Функция	Описание
---------	----------

- | | |
|-------------|-------------------------------------|
| 1) mean() | A) Возвращает минимальное значение |
| 2) min() | B) Вычисляет среднее арифметическое |
| 3) sum() | C) Суммирует элементы вектора |
| 4) length() | D) Определяет количество элементов |

Правильный ответ:

Задание 12

Инструкция: Сопоставьте тип данных в R с примером значения.

Вопрос:

Тип данных	Пример значения
------------	-----------------

- | | |
|--------------|------------|
| 1) numeric | A) "Hello" |
| 2) character | B) TRUE |
| 3) logical | C) 3.14 |

Тип данных Пример значения

4) factor D) factor(c("Low", "High"))

Правильный ответ:

Задание 13

Инструкция: Сопоставьте оператор с его действием.

Вопрос:

Оператор Действие

- 1) == A) Присваивание значения
- 2) != B) Сравнение на равенство
- 3) <- C) Сравнение на неравенство
- 4) %in% D) Проверка вхождения в вектор

Правильный ответ:

Задание 14

Инструкция: Сопоставьте функцию для работы с датафреймами и её назначение.

Вопрос:

Функция Назначение

- 1) head() A) Фильтрация строк по условию
- 2) subset() B) Вывод первых строк датафрейма
- 3) str() C) Отображение структуры данных
- 4) merge() D) Объединение датафреймов

Правильный ответ:

Задание 15

Инструкция: Сопоставьте график с его описанием.

Вопрос:

Тип графика

Описание

- 1) Гистограмма A) Связь между двумя переменными
- 2) Boxplot B) Распределение категориальных данных
- 3) Диаграмма рассеяния C) Распределение количественных данных

Тип графика	Описание
4) Столбчатая диаграмма	D) Разброс и квантили данных

Правильный ответ:

Задание 16

Инструкция: Сопоставьте пакет R с его основной функцией.

Вопрос:

Пакет	Функция
1) dplyr	A) Визуализация данных
2) ggplot2	B) Манипуляции с датафреймами
3) tidyr	C) Преобразование "длинных" данных
4) plotly	D) Создание интерактивных графиков

Правильный ответ:

Задание 17

Инструкция: Сопоставьте функцию для обработки данных с её примером использования.

Вопрос:

Функция	Пример использования
1) filter()	A) select(df, col1, col2)
2) mutate()	B) filter(df, col1 > 10)
3) select()	C) mutate(df, new_col = col1 * 2)
4) group_by()	D) group_by(df, category)

Правильный ответ:

Задание 18

Инструкция: Сопоставьте метод обработки пропущенных значений с его описанием.

Вопрос:

Метод	Описание
1) na.omit()	A) Замена пропусков медианой
2) is.na()	B) Удаление строк с пропусками
3) mean(x, na.rm=TRUE)	C) Проверка на пропуски
4) replace_na()	D) Игнорирование пропусков в расчётах

Правильный ответ:

Задание 19

Инструкция: Сопоставьте семейство функций apply с их назначением.
Вопрос:

Функция	Назначение
---------	------------

- | | |
|-------------|---|
| 1) lapply() | A) Применение к строкам матрицы |
| 2) sapply() | B) Применение к столбцам матрицы |
| 3) apply() | C) Применение к списку, возвращает список |
| 4) vapply() | D) Упрощённый lapply с предопределённым типом |

Правильный ответ:

Задание 20

Инструкция: Сопоставьте термин из анализа данных с его определением.
Вопрос:

Термин	Определение
--------	-------------

- | | |
|--------------|--|
| 1) Outlier | A) Переменная, влияющая на результат |
| 2) Feature | B) Выброс в данных |
| 3) Pipeline | C) Автоматизированная последовательность шагов |
| 4) Reshaping | D) Преобразование структуры данных (напр., из "широкой" в "длинную") |

Правильный ответ:

Задания комбинированного типа (повышенный уровень сложности)

Задание 21

Инструкция: Выберите **все** верные утверждения о работе с пропущенными значениями (NA) в R.
Вопрос:

1. Функция na.omit() удаляет все строки, где хотя бы одно значение — NA.
2. mean(x, na.rm = TRUE) вернет NA, если в векторе x есть пропуски.
3. is.na() применяется только к векторам, но не к датафреймам.
4. replace_na() из пакета tidyr заменяет пропуски указанным значением.

Правильный ответ:

Обоснование:

Задание 22

Инструкция: Выберите **один** вариант, где код приведет к ошибке.
Вопрос:

1. `df %>% filter(!is.na(column))`
2. `df %>% mutate(new_col = column * 2)` (в column есть NA).
3. `df %>% group_by(NA)`
4. `df %>% summarise(mean = mean(column, na.rm = TRUE))`

Правильный ответ:

Обоснование:

Задание 23

Инструкция: Выберите **все** корректные способы выбрать строки датафрейма df, где значение в колонке score больше 5.

Вопрос:

1. `df[df$score > 5,]`
2. `subset(df, score > 5)`
3. `filter(df, score > 5)`
4. `df %>% select(score > 5)`

Правильный ответ:

Обоснование:

Задание 24

Инструкция: Выберите **все** верные утверждения о функциях семейства apply.

Вопрос:

1. `apply()` может работать только с матрицами.
2. `lapply()` всегда возвращает список.
3. `sapply()` пытается упростить результат до вектора или матрицы.
4. `vapply()` безопаснее `sapply()`, так как требует указания типа возвращаемого значения.

Правильный ответ:

Обоснование:

Задание 25

Инструкция: Выберите **один** вариант, где код создаст **интерактивный** график.

Вопрос:

1. `ggplot(mtcars, aes(mpg, hp)) + geom_point()`
2. `plot_ly(mtcars, x = ~mpg, y = ~hp, type = "scatter")`
3. `hist(mtcars$mpg)`
4. `boxplot(mpg ~ cyl, data = mtcars)`

Правильный ответ:

Обоснование:

Задание 26

Инструкция: Выберите **все** корректные способы загрузить данные из Excel-файла в R.

Вопрос:

1. `read.csv("data.xlsx")`
2. `read_excel("data.xlsx")` (из пакета readxl).
3. `load("data.xlsx")`
4. `openxlsx::read.xlsx("data.xlsx")`

Правильный ответ:

Обоснование:

Задание 27

Инструкция: Выберите **все** верные утверждения о dplyr.

Вопрос:

1. mutate() изменяет существующие колонки, но не создает новые.
2. group_by() можно комбинировать с summarise() для агрегации данных.
3. rename() меняет названия колонок без изменения данных.
4. arrange() сортирует строки по убыванию только с помощью desc().

Правильный ответ:

Обоснование:

Задание 28

Инструкция: Выберите **один** вариант, где код **не** создаст фрейм данных.

Вопрос:

1. data.frame(x = 1:3, y = c("a", "b", "c"))
2. tibble(x = 1:3, y = c("a", "b", "c"))
3. list(x = 1:3, y = c("a", "b", "c"))
4. as.data.frame(list(x = 1:3, y = c("a", "b", "c")))

Правильный ответ:

Обоснование:

Задание 29

Инструкция: Выберите **все** функции, которые могут обрабатывать **текстовые данные** в R.

Вопрос:

1. gsub()
2. str_detect() (из stringr).
3. paste()
4. round()

Правильный ответ:

Обоснование:

Задание 30

Инструкция: Выберите **все** верные утверждения о ggplot2.

Вопрос:

1. aes() задает только оси графика, но не цвет или размер.
2. geom_point() требует обязательного указания x и y.
3. theme() изменяет только заголовок графика.
4. facet_wrap() создает множественные графики по категориям.

Правильный ответ:

Обоснование:

Задания открытого типа (высокий уровень сложности)

Задание 31

Инструкция: Дайте развернутый ответ.

Вопрос:

Какие преимущества дает использование dplyr по сравнению с базовым R для обработки данных? Приведите 3 конкретных примера функций из dplyr и их аналогов в базовом R, объяснив разницу.

Поле для ответа:

Ответ-эталон:

Примечание:

Задание 32

Инструкция: Дайте развернутый ответ.

Вопрос:

Опишите, как можно обнаружить и обработать выбросы в данных средствами R.

Упомяните не менее двух методов обнаружения (например, статистических и визуальных) и два способа обработки (удаление, преобразование).

Поле для ответа:

Ответ-эталон:

Примечание:

Задание 33

Инструкция: Дайте развернутый ответ.

Вопрос:

Какие проблемы могут возникнуть при слиянии датафреймов в R и как их избежать?

Укажите не менее двух типичных ошибок (например, разные названия колонок) и решения для каждой.

Поле для ответа:

Ответ-эталон:

Примечание:

Задание 34

Инструкция: Дайте развернутый ответ.

Вопрос:

Объясните, зачем нужны факторы (factor) в R и в каких случаях их использование обязательно. Приведите примеры функций, которые требуют факторных переменных.

Поле для ответа:

Ответ-эталон:

Примечание:

Задание 35

Инструкция: Дайте развернутый ответ.

Вопрос:

Как обеспечить воспроизводимость (reproducibility) анализа в R? Назовите не менее трех инструментов или практик и поясните их роль.

Поле для ответа:

Ответ-эталон:

Примечание:

Задание 36

Инструкция: Дайте развернутый ответ.

Вопрос:

Какие преимущества и недостатки есть у data.table по сравнению с dplyr? Укажите не менее двух пунктов для каждого пакета.

Поле для ответа:

Ответ-эталон:

Примечание:

Задание 37

Инструкция: Дайте развернутый ответ.

Вопрос:

Поле для ответа:

Ответ-эталон:

Примечание:

Задание 38

Инструкция: Дайте развернутый ответ.

Вопрос:

Когда следует использовать списки (list) в R, а когда — векторы? Приведите по два примера задач для каждой структуры.

Поле для ответа:

Ответ-эталон:

Примечание:

Задание 39

Инструкция: Дайте развернутый ответ.

Вопрос:

Какие есть способы обработки текстовых данных в R? Назовите два пакета (например, stringr) и по одной функции из каждого с примерами их использования.

Поле для ответа:

Ответ-эталон:

Примечание:

Задание 40

Инструкция: Дайте развернутый ответ.

Вопрос:

Как можно визуализировать временные ряды в R? Укажите два типа графиков (например, lineplot) и пакеты для их построения.

Поле для ответа:

Ответ-эталон:

Примечание: Достаточно двух типов графиков и упоминания пакетов.

Задание 41

Инструкция: Дайте развернутый ответ.

Вопрос:

Объясните, как работает механизм «ленивых вычислений» (lazy evaluation) в R.

Приведите пример, где это поведение может привести к неочевидным ошибкам, и как их избежать.

Поле для ответа:

Ответ-эталон:

Примечание:

Задание 42

Инструкция: Дайте развернутый ответ.

Вопрос:

Какие способы отладки (debugging) кода в R вы знаете? Опишите два метода (например, browser(), traceback()) и случаи, когда они наиболее полезны.

Поле для ответа:

Ответ-эталон:

Примечание:

Задание 43

Инструкция: Дайте развернутый ответ.

Вопрос:

Как можно оптимизировать потребление памяти при работе с большими датафреймами в R? Назовите два метода (например, использование `data.table`) и объясните их преимущества.

Поле для ответа:

Ответ-эталон:

Примечание:

Задание 44

Инструкция: Дайте развернутый ответ.

Вопрос:

В чем разница между `<-` и `=` в R? Приведите примеры, где использование одного из операторов приведет к ошибке.

Поле для ответа:

Ответ-эталон:

Примечание:

Задание 45

Инструкция: Дайте развернутый ответ.

Вопрос:

Какие есть способы обработки JSON-данных в R? Сравните два пакета по скорости и удобству использования.

Поле для ответа:

Ответ-эталон:

Примечание:

Задание 46

Инструкция: Дайте развернутый ответ.

Вопрос:

Как создать S3-класс в R? Опишите шаги и приведите пример метода для этого класса.

Поле для ответа:

Ответ-эталон:

Примечание:

Задание 47

Инструкция: Дайте развернутый ответ.

Вопрос:

Какие есть альтернативы циклу `for` для применения функции к каждому элементу списка? Сравните `lapply()`, `purrr::map()` и `parallel::mclapply()` по производительности и синтаксису.

Поле для ответа:

Ответ-эталон:

Примечание:

Задание 48

Инструкция: Дайте развернутый ответ.

Вопрос:

Как можно динамически генерировать код в R? Опишите использование `eval()`, `parse()` и `do.call()` с примерами.

Поле для ответа:

Ответ-эталон:

Примечание:

Задание 49

Инструкция: Дайте развернутый ответ.

Вопрос:

Какие есть способы взаимодействия R с другими языками (Python, C++)? Опишите один метод и его ограничения.

Поле для ответа:

Ответ-эталон:

Примечание:

Задание 50

Инструкция: Дайте развернутый ответ.

Вопрос:

Как реализовать кэширование результатов функции в R? Сравните memoise и ручное кэширование через глобальные переменные.

Поле для ответа:

Ответ-эталон:

Примечание:

ОПК-4 (ПИ) Способен применять на практике новые научные принципы и методы исследований

Задания закрытого типа (базовый уровень сложности)

Задание 51

Инструкция: Выберите один правильный ответ.

Вопрос: Какой метод в R позволяет воспроизводить случайные числа при каждом запуске кода?

Варианты ответа:

1. runif()
2. set.seed()
3. sample()
4. rnorm()

Правильный ответ:

Задание 52

Инструкция: Выберите один правильный ответ.

Вопрос: Какой пакет в R используется для создания интерактивных веб-приложений без знания JavaScript?

Варианты ответа:

1. ggplot2
2. shiny
3. plotly
4. dplyr

Правильный ответ:

Задание 53

Инструкция: Выберите один правильный ответ.

Вопрос: Какой функцией в R можно создать датафрейм из списка с разной длиной элементов, автоматически заполнив недостающие значения?

Варианты ответа:

1. data.frame()

2. `as.data.frame()`
3. `tibble()` (из пакета `tibble`)
4. `cbind()`

Правильный ответ:

Задание 54

Инструкция: Выберите один правильный ответ.

Вопрос: Какой метод используется для чтения данных из SQL-базы непосредственно в R?

Варианты ответа:

1. `read.csv()`
2. `DBI::dbGetQuery()`
3. `read_excel()`
4. `jsonlite::fromJSON()`

Правильный ответ:

Задание 55

Инструкция: Выберите один правильный ответ.

Вопрос: Какой оператор в R позволяет временно изменять глобальные настройки (например, локализацию или параметры графиков)?

Варианты ответа:

1. `with()`
2. `<<-`
3. `options()`
4. `par()`

Правильный ответ:

Задание 56

Инструкция: Выберите один правильный ответ.

Вопрос: Какой пакет предоставляет функции для работы с временными рядами, включая декомпозицию и прогнозирование?

Варианты ответа:

1. `dplyr`
2. `forecast`
3. `ggplot2`
4. `tidyr`

Правильный ответ:

Задание 57

Инструкция: Выберите один правильный ответ.

Вопрос: Какой функцией можно преобразовать "широкий" датафрейм в "длинный" формат?

Варианты ответа:

1. `spread()`
2. `gather()`
3. `separate()`
4. `unite()`

Правильный ответ:

Задание 58

Инструкция: Выберите один правильный ответ.

Вопрос: Какой метод в R позволяет автоматически генерировать отчеты с кодом,

графиками и текстом в одном документе?

Варианты ответа:

1. R Markdown
2. shiny
3. ggplot2
4. dplyr

Правильный ответ:

Задание 59

Инструкция: Выберите один правильный ответ.

Вопрос: Какой функцией можно проверить, установлен ли пакет в R, перед его загрузкой?

Варианты ответа:

1. `install.packages()`
2. `library()`
3. `require()`
4. `"package_name" %in% installed.packages()`

Правильный ответ:

Задание 60

Инструкция: Выберите один правильный ответ.

Вопрос: Какой пакет используется для работы с геопространственными данными и построения карт в R?

Варианты ответа:

1. ggplot2
2. sf
3. dplyr
4. tidy

Правильный ответ:

Задания комбинированного типа (повышенный уровень сложности)

Задание 61

Инструкция: Выберите **все** верные утверждения о работе с большими данными в R.

Вопрос:

1. Пакет `data.table` всегда эффективнее `dplyr` при обработке датафреймов размером менее 10 000 строк.
2. Функция `fread()` из `data.table` для чтения CSV-файлов быстрее `read.csv()`, но потребляет больше памяти.
3. Использование `ff` или `bigmemory` позволяет работать с данными, не помещающимися в оперативную память.
4. `dplyr` можно использовать для обработки данных размером больше 100 Гб без дополнительных оптимизаций.

Правильный ответ:

Обоснование:

Задание 62

Инструкция: Выберите **один** вариант, где код точно приведет к ошибке.

Вопрос:

1. `df %>% mutate(new_col = NA)`
2. `df %>% filter(column == "text")` (в `column` есть `NA`).
3. `df %>% group_by(NA)`
4. `df %>% summarise(mean = mean(column, na.rm = TRUE))`

Правильный ответ:

Обоснование:

Задание 63

Инструкция: Выберите **все** верные утверждения о функциональном программировании в R.

Вопрос:

1. `purrr::map()` всегда возвращает список, даже если входной вектор — числовой.
2. `Reduce()` применяет функцию последовательно к элементам списка, накапливая результат.
3. Замыкания (closures) в R не могут изменять состояние внешних переменных.
4. `memoise::memoise()` кэширует результаты функции на основе её тела, а не аргументов.

Правильный ответ:

Обоснование:

Задание 64

Инструкция: Выберите **все** корректные способы добавить новую колонку в датафрейм `df` на основе условия.

Вопрос:

1. `df$new_col <- ifelse(df$old_col > 5, "A", "B")`
2. `df %>% mutate(new_col = case_when(old_col > 5 ~ "A", TRUE ~ "B"))`
3. `df[, "new_col"] <- lapply(df$old_col, function(x) if (x > 5) "A" else "B")`
4. `transform(df, new_col = if (old_col > 5) "A" else "B")`

Правильный ответ:

Обоснование:

Задание 65

Инструкция: Выберите **один** вариант, где код создаст **интерактивную** таблицу.

Вопрос:

1. `knitr::kable(df)`
2. `DT::datatable(df)`
3. `ggplot2::ggplot(df)`
4. `dplyr::glimpse(df)`

Правильный ответ:

Обоснование:

Задание 66

Инструкция: Выберите **все** верные утверждения о параллельных вычислениях в R.

Вопрос:

1. Пакет `parallel` не работает на Windows.
2. `future::plan()` позволяет выбрать режим параллелизации (например, многопоточность или кластер).
3. `foreach()` с `%dopar%` требует регистрации бэкенда (например, `doParallel`).
4. Параллельные вычисления всегда ускоряют код в 10+ раз.

Правильный ответ:

Обоснование:

Задание 67

Инструкция: Выберите **все** корректные способы удалить дубликаты строк в датафрейме `df`.

Вопрос:

1. `df[!duplicated(df),]`

2. `df %>% distinct()`
3. `unique(df)`
4. `df %>% group_by_all() %>% slice(1)`

Правильный ответ:

Обоснование:

Задание 68

Инструкция: Выберите **один** вариант, где код **не** будет работать с типом `tibble`.

Вопрос:

1. `df[, 1]`
2. `df[[1]]`
3. `df$column`
4. `df[1, "column"]`

Правильный ответ:

Обоснование:

Индексация `df[, 1]` (1) вернет `tibble`, а не вектор, что может нарушить ожидания в коде, рассчитанном на `data.frame`. Остальные варианты (`df[[1]]`, `df$`, `df[i, j]`) работают идентично для `tibble` и `data.frame`.

Задание 69

Инструкция: Выберите **все** верные утверждения о S3-классах в R.

Вопрос:

1. Методы для S3-классов создаются через `function.class_name()`.
2. `print()` автоматически вызывает метод `print.class_name()`, если объект имеет класс `class_name`.
3. S3 не поддерживает наследование.
4. Проверить класс объекта можно только через `class()`.

Правильный ответ:

Обоснование:

Задание 70

Инструкция: Выберите **все** корректные способы сохранить датафрейм `df` в файл.

Вопрос:

1. `write.csv(df, "file.csv")`
2. `saveRDS(df, "file.rds")`
3. `write.table(df, "file.txt")`
4. `jsonlite::write_json(df, "file.json")`

Правильный ответ:

Обоснование:

Задания закрытого типа (повышенный уровень сложности)

Задание 71

Инструкция: Сопоставьте функцию из `dplyr` с её аналогом в базовом R.

dplyr

База R

- | | |
|-----------------------------|-----------------------------|
| 1) <code>filter()</code> | A) <code>subset()</code> |
| 2) <code>mutate()</code> | B) <code>transform()</code> |
| 3) <code>summarise()</code> | C) <code>aggregate()</code> |

dplyr База R

4) arrange() D) order()

Правильный ответ:

Задание 72

Инструкция: Сопоставьте пакет с его специализацией.

Пакет	Применение
--------------	-------------------

1) lubridate	A) Работа с геопространственными данными
--------------	--

2) sf	B) Обработка дат и времени
-------	----------------------------

3) stringr	C) Манипуляции с текстом
------------	--------------------------

4) purrr	D) Функциональное программирование
----------	------------------------------------

Правильный ответ:

Задание 73

Инструкция: Сопоставьте тип ошибки в R с примером кода, который её вызывает.

Ошибка	Пример кода
---------------	--------------------

1) NULL-обращение	A) x <- list(a = 1); x\$b
-------------------	---------------------------

2) Индексация NA	B) df[NA,]
------------------	-------------

3) Деление на ноль	C) sum("text")
--------------------	----------------

4) Несовместимость типов	D) 1 / 0
--------------------------	----------

Правильный ответ:

Задание 74

Инструкция: Сопоставьте метод оптимизации кода с примером его использования.

Метод	Пример
--------------	---------------

1) Векторизация	A) Замена цикла for на apply()
-----------------	--------------------------------

2) Мемоизация	B) memoise::memoise(factorial)
---------------	--------------------------------

3) Параллелизация	C) parallel::mclapply(1:10, fun)
-------------------	----------------------------------

4) Профилирование	D) profvis::profvis({ код })
-------------------	------------------------------

Правильный ответ:

Задание 75

Инструкция: Сопоставьте функцию для работы с датами и её результат.

Функция	Результат
1) lubridate::ymd()	A) 2023-01-01 из строки "20230101"
2) as.POSIXct()	B) Разница между датами в днях
3) difftime()	C) Текущая дата и время
4) Sys.time()	D) Преобразует строку в timestamp

Правильный ответ:

Задание 76

Инструкция: Сопоставьте задачу с подходящим пакетом для её решения.

Задача	Пакет
1) Визуализация сетевых графов	A) igraph
2) Парсинг HTML-страниц	B) rvest
3) Работа с матрицами	C) Matrix
4) Анализ временных рядов	D) forecast

Правильный ответ:

Задание 77

Инструкция: Сопоставьте оператор с его действием в dplyr.

Оператор	Действие
1) %>%	A) Фильтрация строк
2) %<>%	B) Конвейер с перезаписью
3) %\$%	C) Доступ к колонкам как к переменным
4) %in%	D) Проверка вхождения в вектор

Правильный ответ:

Задание 78

Инструкция: Сопоставьте функцию для работы с текстом и её применение.

Функция	Применение
1) str_detect()	A) Замена подстрок по шаблону
2) str_replace()	B) Разделение строки по разделителю
3) str_split()	C) Поиск совпадений с регулярным выражением

Функция	Применение
---------	------------

4) str_trim()	D) Удаление пробелов по краям
---------------	-------------------------------

Правильный ответ:

Задание 79

Инструкция: Сопоставьте метод отладки с ситуацией, где он полезен.

Метод	Ситуация
-------	----------

- | | |
|----------------|--|
| 1) browser() | A) Пошаговый анализ функции |
| 2) traceback() | B) Поиск места ошибки в стеке вызовов |
| 3) debug() | C) Автоматическая отладка функции |
| 4) tryCatch() | D) Обработка ошибок без остановки кода |

Правильный ответ:

Задание 80

Инструкция: Сопоставьте структуру данных с её описанием.

Структура	Описание
-----------	----------

- | | |
|------------|--|
| 1) Список | A) Одномерный массив одного типа |
| 2) Вектор | B) Упорядоченная коллекция разнотипных элементов |
| 3) Матрица | C) Двумерный массив одного типа |
| 4) Фрейм | D) Таблица с колонками разных типов |

Правильный ответ:

Задания открытого типа (высокий уровень сложности)

Задание 81

Инструкция: Дайте развернутый ответ.

Вопрос:

Опишите стратегию обработки данных, где строки содержат частично дублирующуюся информацию (например, повторяющиеся клиенты с разными датами заказов). Какие функции R вы бы использовали для выявления и устранения таких случаев?

Поле для ответа:

Ответ-эталон:

Примечание:

Задание 82

Инструкция: Дайте развернутый ответ.

Вопрос:

Как организовать мониторинг выполнения длительного R-скрипта (например, обработки большого файла), чтобы отслеживать прогресс и возможные ошибки без постоянного ручного контроля?

Поле для ответа:

Ответ-эталон:

Примечание:

Задание 83

Инструкция: Дайте

развернутый

ответ.

Вопрос:

Какие методы применяются для безопасной обработки конфиденциальных данных (например, персональных идентификаторов) в R, чтобы избежать их случайного сохранения или логирования?

Поле для ответа:

Ответ-эталон:

Примечание:

Задание 84

Инструкция: Дайте

развернутый

ответ.

Вопрос:

Как создать динамический пользовательский интерфейс в Shiny, где элементы (например, фильтры или графики) генерируются на основе выбора пользователя?

Поле для ответа:

Ответ-эталон:

Примечание:

Задание 85

Инструкция: Дайте

развернутый

ответ.

Вопрос:

Какие подходы в R позволяют эффективно работать с данными, обновляемыми в реальном времени (например, биржевые котировки)?

Поле для ответа:

Ответ-эталон:

Примечание:

Задание 86

Инструкция: Дайте

развернутый

ответ.

Вопрос:

Опишите, как реализовать многопоточность в R для ускорения обработки данных, и какие ограничения следует учитывать?

Поле для ответа:

Ответ-эталон:

Примечание:

Задание 87

Инструкция: Дайте

развернутый

ответ.

Вопрос:

Как организовать тестирование R-пакета, чтобы проверить корректность функций и обработать крайние случаи?

Поле для ответа:

Ответ-эталон:

Примечание:

Задание 88

Инструкция: Дайте развернутый ответ.

Вопрос:

Какие инструменты R помогают документировать код и автоматизировать создание справочных материалов для пользователей?

Поле для ответа:

Ответ-эталон:

Примечание:

Задание 89

Инструкция: Дайте развернутый ответ.

Вопрос:

Как создать пользовательский S3-класс в R для представления специализированных данных (например, геокоординат) и перегрузить для него стандартные операции?

Поле для ответа:

Ответ-эталон:

Примечание:

Задание 90

Инструкция: Дайте развернутый ответ.

Вопрос:

Какие методы в R позволяют работать с данными, размер которых превышает объем оперативной памяти?

Поле для ответа:

Ответ-эталон:

Примечание:

Задание 91

Инструкция: Дайте развернутый ответ.

Вопрос:

Как реализовать в R пользовательскую функцию, которая принимает произвольное количество аргументов разного типа (числа, строки, списки) и возвращает их в виде структурированного списка с сохранением типов?

Поле для ответа:

Ответ-эталон:

Примечание:

Задание 92

Инструкция: Дайте развернутый ответ.

Вопрос:

Каким образом можно визуализировать многомерные данные (более 3 переменных) на одном графике без потери информативности?

Поле для ответа:

Ответ-эталон:

Примечание:

Задание 93

Инструкция: Дайте развернутый ответ.

Вопрос:

Как создать анимированный график изменения данных во времени (например, динамика цен акций) с использованием R?

Поле для ответа:

Ответ-эталон:

Примечание:

Задание 94

Инструкция: Дайте развернутый ответ.

Вопрос:

Какие методы в R позволяют эффективно сравнивать два датафрейма с одинаковой структурой, но потенциально различающимися значениями?

Поле для ответа:

Ответ-эталон:

Примечание:

Задание 95

Инструкция: Дайте развернутый ответ.

Вопрос:

Как в R реализовать обработку данных по принципу "ленивых вычислений" (lazy evaluation), когда вычисления происходят только при непосредственном обращении к объекту?

Поле для ответа:

Ответ-эталон:

Примечание:

Задание 96

Инструкция: Дайте развернутый ответ.

Вопрос:

Каким образом можно программно определить зависимости между функциями внутри пользовательского R-пакета для построения графа их взаимодействий?

Поле для ответа:

Ответ-эталон:

Примечание:

Задание 97

Инструкция: Дайте развернутый ответ.

Вопрос:

Как создать в R функцию с возможностью "цепочки" (chaining), подобную оператору `%>%`, но с дополнительной логикой (например, логгированием промежуточных результатов)?

Поле для ответа:

Ответ-эталон:

Примечание:

Задание 98

Инструкция: Дайте развернутый ответ.

Вопрос:

Какие существуют способы обработки вложенных JSON-структур (например, API-ответов с переменной глубиной) в R без предварительного знания их схемы?

Поле для ответа:

Ответ-эталон:

Примечание:

Задание 99

Инструкция: Дайте развернутый ответ.

Вопрос:

Как в R реализовать "горячее" обновление Shiny-приложения (без перезагрузки страницы) при изменении исходных данных на сервере?

Поле для ответа:

Ответ-эталон:

Примечание:

Задание 100

Инструкция: Дайте развернутый ответ.

Вопрос:

Какие методы в R позволяют оптимизировать работу с факторными переменными (factors) для уменьшения потребления памяти и ускорения операций?

Поле для ответа:

Ответ-эталон:

Примечание:

5 Средства оценки индикаторов достижения компетенций

Таблица 4

Средства оценки индикаторов достижения компетенций

Коды компетенций	Индикаторы компетенций (в соот.с Таблицей 1)	Средства оценки (в соот. с Таблицами 5, 7)
ОПК-1 (ПИ)	ИД.ОПК-1.1. ИД.ОПК-1.2.	Домашнее задание, тест
ОПК-4 (ПИ)	ИД.ОПК-4.1. ИД.ОПК-4.2. ИД.ОПК-4.3. ИД.ОПК-4.4.	Домашнее задание, тест

Таблица 5

Описание средств оценки индикаторов достижения компетенций

Средства оценки (в соот. с Таблицами 5, 7)	Рекомендованный план выполнения работы
Домашнее задание	<p>Магистрант в ходе подготовки и выполнения домашнего задания показывает наличие практической базы знаний в рамках дисциплины, необходимой для выполнения следующих действий в области профессиональной деятельности:</p> <p>Делать обоснованный выбор методов разработки алгоритмов и программных средств для решения профессиональных задач, в том числе с использованием современных информационных технологий</p> <p>На основе современных теорий и концепций обосновывать актуальность, анализировать, применять и разрабатывать предложения и рекомендации по использованию новых научных принципов и методов исследований в профессиональной области знаний</p>
Тест	<p>Магистрант в ходе подготовки и выполнения тестирования показывает наличие практической базы знаний в рамках дисциплины, необходимой для выполнения следующих действий в области профессиональной деятельности:</p> <p>Делать обоснованный выбор методов разработки алгоритмов и программных средств для решения профессиональных задач, в том числе с использованием современных информационных технологий</p> <p>На основе современных теорий и концепций обосновывать актуальность, анализировать, применять и разрабатывать предложения и рекомендации по</p>

Средства оценки <i>(в соот. с Таблицами 5, 7)</i>	Рекомендованный план выполнения работы
	использованию новых научных принципов и методов исследований в профессиональной области знаний