

Документ подписан простой электронной подписью  
Информация о владельце:

ФИО: Кандрашина Елена Александровна

Должность: Врио ректора ФГАОУ ВО «Самарский государственный экономический университет»

Дата подписания: 07.07.2023 15:32:39

Уникальный программный ключ:

b2fd765521f4c570b8c6e8e502a10b4f1de8ae0d

**Министерство науки и высшего образования Российской Федерации**  
**Федеральное государственное автономное образовательное учреждение**  
**высшего образования**  
**«Самарский государственный экономический университет»**

**Институт**      Институт экономики предприятий

**Кафедра**      Прикладной информатики

**УТВЕРЖДЕНО**

Ученым советом Университета

(протокол № 11 от 30 мая 2023 г.)

**РАБОЧАЯ ПРОГРАММА**

**Наименование дисциплины**      Б1.В.11 Технологии больших данных

**Основная профессиональная образовательная программа**      09.03.03 Прикладная информатика программа  
Цифровые технологии в экономике

Квалификация (степень) выпускника Бакалавр

Самара 2023

## Содержание (рабочая программа)

Стр.

- 1 Место дисциплины в структуре ОП
- 2 Планируемые результаты обучения по дисциплине, обеспечивающие достижение планируемых результатов обучения по программе
- 3 Объем и виды учебной работы
- 4 Содержание дисциплины
- 5 Материально-техническое и учебно-методическое обеспечение дисциплины
- 6 Фонд оценочных средств по дисциплине

Целью изучения дисциплины является формирование результатов обучения, обеспечивающих достижение планируемых результатов освоения образовательной программы.

## 1. Место дисциплины в структуре ОП

Дисциплина Технологии больших данных входит в часть, формируемая участниками образовательных отношений блока Б1. Дисциплины (модули)

Предшествующие дисциплины по связям компетенций: Хранение, обработка и анализ данных, Системы искусственного интеллекта, Методы оптимизации и теория игр, Разработка интерфейсов и адаптивный Веб-дизайн, Интеллектуальные информационные системы, Технологии работы в социальных сетях, Информационно-коммуникационные технологии в профессиональной деятельности, Основы алгоритмизации и программирования, Основы проектной деятельности, Современные технологии и языки программирования, Встроенные языки программирования, Организация вычислительных процессов, Философия, История России, Математические методы в экономике, Общая теория статистики, Основы финансовых расчетов, Эконометрика, Управление человеческими ресурсами, Основы менеджмента, Облачные технологии и услуги

Последующие дисциплины по связям компетенций: Цифровые технологии управления предприятием, Современные цифровые платформы, Разработка профессиональных приложений, Управление ИТ-проектами, Разработка мобильных приложений, Интернет-предпринимательство, Проектный практикум, Управление качеством разработки приложений, Проектирование информационных систем

## 2. Планируемые результаты обучения по дисциплине, обеспечивающие достижение планируемых результатов обучения по программе

Изучение дисциплины Технологии больших данных в образовательной программе направлено на формирование у обучающихся следующих компетенций:

### Универсальные компетенции (УК):

УК-1 - Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач

Планируемые результаты обучения по программе	Планируемые результаты обучения по дисциплине		
УК-1	УК-1.1: Знать:	УК-1.2: Уметь:	УК-1.3: Владеть (иметь навыки):
	методы поиска, анализа и синтеза информации	осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач	навыками поиска, критического анализа и синтеза информации, применения системного подхода для решения поставленных задач

### Профессиональные компетенции (ПК):

ПК-4 - Способен к верификации структуры программного кода ИС относительно архитектуры ИС и требований заказчика к ИС

Планируемые результаты обучения по программе	Планируемые результаты обучения по дисциплине		
ПК-4	ПК-4.1: Знать:	ПК-4.2: Уметь:	ПК-4.3: Владеть (иметь навыки):
	особенности верификации структуры программного кода ИС	верифицировать структуру программного кода ИС относительно	навыками верификации структуры программного кода ИС относительно

	относительно архитектуры ИС и требований заказчика к ИС	архитектуры ИС и требований заказчика к ИС	архитектуры ИС и требований заказчика к ИС
--	---	--	--

ПК-3 - Способен к разработке структуры программного кода ИС

Планируемые результаты обучения по программе	<b>Планируемые результаты обучения по дисциплине</b>		
ПК-3	ПК-3.1: Знать:	ПК-3.2: Уметь:	ПК-3.3: Владеть (иметь навыки):
	особенности разработки структуры программного кода ИС	разрабатывать структуру программного кода ИС	навыками разработки структуры программного кода ИС

### 3. Объем и виды учебной работы

Учебным планом предусматриваются следующие виды учебной работы по дисциплине:

#### Очная форма обучения

Виды учебной работы	Всего час/ з.е.
	Сем 6
Контактная работа, в том числе:	56.3/1.56
Занятия лекционного типа	18/0.5
Лабораторные работы (лабораторный практикум)	36/1
Индивидуальная контактная работа (ИКР)	0.3/0.01
Групповая контактная работа (ГКР)	2/0.06
Самостоятельная работа:	53.7/1.49
Промежуточная аттестация	34/0.94
Вид промежуточной аттестации:	
Экзамен	Экз
Общая трудоемкость (объем части образовательной программы): Часы	144
Зачетные единицы	4

#### заочная форма

Виды учебной работы	Всего час/ з.е.
	Сем 6
Контактная работа, в том числе:	6.3/0.18
Занятия лекционного типа	2/0.06
Лабораторные работы (лабораторный практикум)	2/0.06
Индивидуальная контактная работа (ИКР)	0.3/0.01
Групповая контактная работа (ГКР)	2/0.06
Самостоятельная работа:	103.7/2.88
Промежуточная аттестация	34/0.94
Вид промежуточной аттестации:	
Экзамен	Экз
Общая трудоемкость (объем части образовательной программы): Часы	144
Зачетные единицы	4

### 4. Содержание дисциплины

#### 4.1. Разделы, темы дисциплины и виды занятий:

Тематический план дисциплины Технологии больших данных представлен в таблице.

**Разделы, темы дисциплины и виды занятий  
Очная форма обучения**

№ п/п	Наименование темы (раздела) дисциплины	Контактная работа				Самостоятельная работа	Планируемые результаты обучения в соотношении с результатами обучения по образовательной программе
		Лекции	Занятия семинарского типа	ИКР	ГКР		
			Лаборат. работы				
1.	Современные инструменты и алгоритмы анализа BIG DATA.	9	18	0,15	1	26,85	УК-1.1, УК-1.2, УК-1.3, ПК-4.1, ПК-4.2, ПК-4.3, ПК- 3.1, ПК-3.2, ПК-3.3
2.	Язык R один из стандартов при работе с данными и статистическая обработка данных.	9	18	0,15	1	26,85	УК-1.1, УК-1.2, УК-1.3, ПК-4.1, ПК-4.2, ПК-4.3, ПК- 3.1, ПК-3.2, ПК-3.3
Контроль		34					
<b>Итого</b>		<b>18</b>	<b>36</b>	<b>0.3</b>	<b>2</b>	<b>53.7</b>	

**заочная форма**

№ п/п	Наименование темы (раздела) дисциплины	Контактная работа				Самостоятельная работа	Планируемые результаты обучения в соотношении с результатами обучения по образовательной программе
		Лекции	Занятия семинарского типа	ИКР	ГКР		
			Лаборат. работы				
1.	Современные инструменты и алгоритмы анализа BIG DATA.	1	1	0,15		51,85	УК-1.1, УК-1.2, УК-1.3, ПК-4.1, ПК-4.2, ПК-4.3, ПК- 3.1, ПК-3.2, ПК-3.3
2.	Язык R один из стандартов при работе с данными и статистическая обработка данных.	1	1	0,15		51,85	УК-1.1, УК-1.2, УК-1.3, ПК-4.1, ПК-4.2, ПК-4.3, ПК- 3.1, ПК-3.2, ПК-3.3
Контроль		34					
<b>Итого</b>		<b>2</b>	<b>2</b>	<b>0.3</b>	<b>2</b>	<b>103.7</b>	

**4.2 Содержание разделов и тем**

**4.2.1 Контактная работа**

**Тематика занятий лекционного типа**

№п/п	Наименование темы (раздела) дисциплины	Вид занятия лекционного типа*	Тематика занятия лекционного типа
1.	Современные инструменты и	лекция	Основные понятия BIG DATA. Технологии и инструментарий анализа больших данных.

	алгоритмы анализа BIG DATA.	лекция	R и RStudio. Математические операции. Векторы и операции с ними. Работа с данными. Типы данных. Датафреймы и операции с ними. Загрузка внешних данных в R.
		лекция	Python. Математика для анализа данных. Работа с матрицами, векторами.
		лекция	Python. Библиотеки для работы с данными. Расширенная визуализация с matplotlib.
2.	Язык R один из стандартов при работе с данными и статистическая обработка данных.	лекция	Визуализация данных и результатов как инструмент.
		лекция	Описательная статистика.
		лекция	Тестирование гипотез. Одновыборочный t-тест. Значение p (p-value) - частые заблуждения. Тестирование гипотез. Двухвыборочный t-тест.
		лекция	Множественное тестирование гипотез.
		лекция	Ошибки при тестировании. Коррекция уровня значимости при множественном тестировании в R.

\*лекции и иные учебные занятия, предусматривающие преимущественную передачу учебной информации педагогическими работниками организации и (или) лицами, привлекаемыми организацией к реализации образовательных программ на иных условиях, обучающимся

#### Тематика занятий семинарского типа

№п/п	Наименование темы (раздела) дисциплины	Вид занятия семинарского типа**	Тематика занятия семинарского типа
1.	Современные инструменты и алгоритмы анализа BIG DATA.	лабораторные работы	Язык R. Настройка рабочего пространства, настройка библиотек.
		лабораторные работы	Язык R. Типы данных, условия, ввод и вывод.
		лабораторные работы	Язык Python . Работа в Анаконде. Основные понятия математического анализа.
		лабораторные работы	Язык Python . Работа в Анаконде. Векторы и матрицы.
		лабораторные работы	Язык Python . Работа в Анаконде. Библиотеки для работы с данными.
		лабораторные работы	Язык Python . Работа в Анаконде. Анализ данных.
		лабораторные работы	Язык R. Операции с векторами.
		лабораторные работы	Язык R. Операции с матрицами, массивами.
2.	Язык R один из стандартов при работе с данными и статистическая обработка данных.	лабораторные работы	Язык R. Описание структуры данных с помощью функций. Циклы.
		лабораторные работы	Язык R. Работа с табличными данными. Структура, чтение данных из внешнего файла.
		лабораторные работы	Язык R. Функциональное программирование

	лабораторные работы	Язык R. Базовая графика. Стандартные графики.
	лабораторные работы	Язык R. Настройка графических средств.
	лабораторные работы	Язык R. Разметка осей, рамка, сетка координат и произвольные линии
	лабораторные работы	Язык R. Описательная статистика. Решение практических задач
	лабораторные работы	Язык R. Тестирование гипотез
	лабораторные работы	Язык R. Коррекция уровня значимости при множественном тестировании в R.

\*\* семинары, практические занятия, практикумы, лабораторные работы, коллоквиумы и иные аналогичные занятия

### Иная контактная работа

При проведении учебных занятий СГЭУ обеспечивает развитие у обучающихся навыков командной работы, межличностной коммуникации, принятия решений, лидерских качеств (включая при необходимости проведение интерактивных лекций, групповых дискуссий, ролевых игр, тренингов, анализ ситуаций и имитационных моделей, преподавание дисциплин (модулей) в форме курсов, составленных на основе результатов научных исследований, проводимых организацией, в том числе с учетом региональных особенностей профессиональной деятельности выпускников и потребностей работодателей).

Формы и методы проведения иной контактной работы приведены в Методических указаниях по основной профессиональной образовательной программе.

#### 4.2.2 Самостоятельная работа

№п/п	Наименование темы (раздела) дисциплины	Вид самостоятельной работы ***
1.	Современные инструменты и алгоритмы анализа BIG DATA.	- подготовка доклада - подготовка электронной презентации - тестирование
2.	Язык R один из стандартов при работе с данными и статистическая обработка данных.	- подготовка доклада - подготовка электронной презентации - тестирование

\*\*\* самостоятельная работа в семестре, написание курсовых работ, докладов, выполнение контрольных работ

## 5. Материально-техническое и учебно-методическое обеспечение дисциплины

### 5.1 Литература:

#### Основная литература

1. Парфенов, Ю. П. Постреляционные хранилища данных : учебное пособие для вузов / Ю. П. Парфенов ; под научной редакцией Н. В. Папуловской. — Москва : Издательство Юрайт, 2023. — 121 с. — (Высшее образование). — ISBN 978-5-534-09837-2. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/514724>

#### Дополнительная литература

1. Гордеев, С. И. Организация баз данных в 2 ч. Часть 1 : учебник для вузов / С. И. Гордеев, В. Н. Волошина. — 2-е изд., испр. и доп. — Москва : Издательство Юрайт, 2023. — 310 с. — (Высшее образование). — ISBN 978-5-534-04469-0. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/513879>

2. Гордеев, С. И. Организация баз данных в 2 ч. Часть 2 : учебник для вузов / С. И. Гордеев, В. Н. Волошина. — 2-е изд., испр. и доп. — Москва : Издательство Юрайт, 2023. — 513 с. — (Высшее образование). — ISBN 978-5-534-04470-6. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/515097>

### 5.2. Перечень лицензионного программного обеспечения

1. Microsoft Windows 10 Education / Microsoft Windows 7 / Windows Vista Business
2. Microsoft Office 2016 Professional Plus (Word, Excel, Access, PowerPoint, Outlook, OneNote, Publisher) / Microsoft Office 2007 (Word, Excel, Access, PowerPoint)
2. Анаконда <https://www.anaconda.com/download/> - для Windows/ Linux/ OS
3. RSudio
4. Loginom

### **5.3 Современные профессиональные базы данных, к которым обеспечивается доступ обучающихся**

1. Профессиональная база данных «Информационные системы Министерства экономического развития Российской Федерации в сети Интернет» (Портал «Официальная Россия» - <http://www.gov.ru/>)
2. Профессиональная база данных «Финансово-экономические показатели Российской Федерации» (Официальный сайт Министерства финансов РФ - <https://www.minfin.ru/ru/>)
3. Профессиональная база данных «Официальная статистика» (Официальный сайт Федеральной службы государственной статистики - <http://www.gks.ru/>)

### **5.4. Информационно-справочные системы, к которым обеспечивается доступ обучающихся**

1. Справочно-правовая система «Консультант Плюс»
2. Справочно-правовая система «ГАРАНТ-Максимум»

### **5.5. Специальные помещения**

Учебные аудитории для проведения занятий лекционного типа	Комплекты ученической мебели Мультимедийный проектор Доска Экран
Учебные аудитории для проведения практических занятий (занятий семинарского типа)	Комплекты ученической мебели Мультимедийный проектор Доска Экран Компьютеры с выходом в сеть «Интернет» и ЭИОС СГЭУ
Учебные аудитории для групповых и индивидуальных консультаций	Комплекты ученической мебели Мультимедийный проектор Доска Экран Компьютеры с выходом в сеть «Интернет» и ЭИОС СГЭУ
Учебные аудитории для текущего контроля и промежуточной аттестации	Комплекты ученической мебели Мультимедийный проектор Доска Экран Компьютеры с выходом в сеть «Интернет» и ЭИОС СГЭУ
Помещения для самостоятельной работы	Комплекты ученической мебели Мультимедийный проектор Доска Экран Компьютеры с выходом в сеть «Интернет» и ЭИОС СГЭУ
Помещения для хранения и профилактического обслуживания оборудования	Комплекты специализированной мебели для хранения оборудования



## 5.6 Лаборатории и лабораторное оборудование

Лаборатория информационных технологий в профессиональной деятельности	Комплекты ученической мебели Мультимедийный проектор Доска Экран Компьютеры с выходом в сеть «Интернет» и ЭИОС СГЭУ Лабораторное оборудование
---	--

## 6. Фонд оценочных средств по дисциплине Технологии больших данных:

### 6.1. Контрольные мероприятия по дисциплине

Вид контроля	Форма контроля	Отметить нужное знаком « + »
Текущий контроль	Тестирование	+
	Лабораторные работы	+
Промежуточный контроль	Экзамен	+

Порядок проведения мероприятий текущего и промежуточного контроля определяется Методическими указаниями по основной профессиональной образовательной программе высшего образования; Положением о балльно-рейтинговой системе оценки успеваемости обучающихся по основным образовательным программам высшего образования – программам бакалавриата, программам специалитета, программам магистратуры в федеральном государственном автономном образовательном учреждении высшего образования «Самарский государственный экономический университет».

### 6.2. Планируемые результаты обучения по дисциплине, обеспечивающие достижение планируемых результатов обучения по программе

#### Универсальные компетенции (УК):

УК-1 - Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач

Планируемые результаты обучения по программе	Планируемые результаты обучения по дисциплине		
	УК-1.1: Знать:	УК-1.2: Уметь:	УК-1.3: Владеть (иметь навыки):
	методы поиска, анализа и синтеза информации	осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач	навыками поиска, критического анализа и синтеза информации, применения системного подхода для решения поставленных задач
Пороговый	аспекты критического анализа информации	определять насколько правдивости, достоверности или вероятности представленной информации	анализа исследований, научных фактов, построение выводов

Стандартный (в дополнение к пороговому)	видах мышления, основы формальной логика	анализировать факты, выявить необоснованные суждения и ложные мнения  анализировать исследования, научные факты	формулировки решения на основе системного подхода для решения поставленных задач
Повышенный (в дополнение к пороговому, стандартному)	возможности обработки собранной информации для решения профессиональных задач.	систематизировать и интерпретировать полученную информацию для решения профессиональных задач.	приемами решения профессиональных задач на основе результатов, полученных в результате анализа и обработки собранной информации.

**Профессиональные компетенции (ПК):**

ПК-4 - Способен к верификации структуры программного кода ИС относительно архитектуры ИС и требований заказчика к ИС

Планируемые результаты обучения по программе	<b>Планируемые результаты обучения по дисциплине</b>		
	ПК-4.1: Знать:	ПК-4.2: Уметь:	ПК-4.3: Владеть (иметь навыки):
	особенности верификации структуры программного кода ИС относительно архитектуры ИС и требований заказчика к ИС	верифицировать структуру программного кода ИС относительно архитектуры ИС и требований заказчика к ИС	навыками верификации структуры программного кода ИС относительно архитектуры ИС и требований заказчика к ИС
Пороговый	Основные этапы планирования проекта	планировать проект на основе его технико-экономического обоснования, оценки рисков в соответствии с полученным планом, потребностями цифровой экономики	составления технико-экономического обоснования проекта
Стандартный (в дополнение к пороговому)	требования к отказоустойчивости облачных систем исходя из бизнес задач	проектирование под отказ (Design for failure)	планирования проекта на основе его технико-экономического обоснования, оценки рисков в соответствии с полученным планом, потребностями цифровой экономики

Повышенный (в дополнение к пороговому, стандартному)	методики расчетов экономического обоснования и рисков	рассчитывать обоснование оценивать риски	работы с проектами на основе оценки рисков в соответствии с полученным планом, потребностями цифровой экономики
---	---	--	---

ПК-3 - Способен к разработке структуры программного кода ИС

Планируемые результаты обучения по программе	<b>Планируемые результаты обучения по дисциплине</b>		
	ПК-3.1: Знать:	ПК-3.2: Уметь:	ПК-3.3: Владеть (иметь навыки):
	особенности разработки структуры программного кода ИС	разрабатывать структуру программного кода ИС	навыками разработки структуры программного кода ИС
Пороговый	методы анализа информации	собирать и анализировать информацию	навыки анализа информации
Стандартный (в дополнение к пороговому)	этапы подготовки данных и основы работы с базами данных	применять технологии отбора, обработки информации	работы с технологиями извлечения информации из различных источников
Повышенный (в дополнение к пороговому, стандартному)	методы работы с данными, области их применения, достоинства и недостатки, основные цифровые технологии и алгоритмы их работы	интерпретации и анализа информации, отбирать технологии работы с информацией в зависимости от класса задач в области данных	работы представления в понятном виде и передачи, в том числе с помощью цифровых технологий Интерне, методами и технологиями обработки и представления информации

**6.3. Паспорт оценочных материалов**

№ п/п	Наименование темы (раздела) дисциплины	Контролируемые планируемые результаты обучения в соотношении с результатами обучения по программе	Вид контроля/используемые оценочные средства	
			Текущий	Промежуточный
1.	Современные инструменты и алгоритмы анализа BIG DATA.	УК-1.1, УК-1.2, УК- 1.3, ПК-4.1, ПК-4.2, ПК-4.3, ПК-3.1, ПК- 3.2, ПК-3.3	Тестирование Лабораторные работы	Экзамен
2.	Язык R один из стандартов при работе с	УК-1.1, УК-1.2, УК- 1.3, ПК-4.1, ПК-4.2, ПК-4.3, ПК-3.1, ПК- 3.2, ПК-3.3	Тестирование Лабораторные работы	Экзамен

данными и статистическая обработка данных.			
--	--	--	--

#### 6.4.Оценочные материалы для текущего контроля

Ссылка на текущую академическую активность, точки текущего контроля для всех оценочных материалов, размещенных в БРСО ЭИОС СГЭУ: <https://lms2.sseu.ru/course/index.php?categoryid=1819>

#### Примерная тематика докладов

Раздел дисциплины	Темы
Современные инструменты и алгоритмы анализа BIG DATA.	<ol style="list-style-type: none"> <li>1. Big data(Большие данные)</li> <li>2. Международные проекты в области Big Data</li> <li>3. Big Data/ Основы Hadoop. Базовый набор компонентов Hadoop.</li> <li>4. Big Data: обработка больших объемов данных</li> <li>5. Большие данные (Big Data). Назначение, применение, перспектива.</li> <li>6. Применение big data в маркетинговых исследованиях потребителей.</li> <li>7. Большие данные - Big Data - Статкомитет СНГ</li> </ol>
Язык R один из стандартов при работе с данными и статистическая обработка данных.	<ol style="list-style-type: none"> <li>1. Статистический анализ данных в связях с общественностью</li> <li>2. RMarkdown, R и ggplot</li> <li>3. Прогнозные модели посещаемости сайта в RStudio</li> <li>4. RStudio, можно будет построить интерактивную диаграмму</li> <li>5. Data Science для карьериста</li> <li>6. Среда разработки RStudio</li> <li>7. Особенности эконом. состояния регионов страны. Статистическая оценка эконом. состояния регионов страны в RStudio.</li> <li>8. Анализ структуры расходов федерального бюджета за 2018 и 2020 год в RStudio.</li> <li>9. Статистика, аналитика, R-studio</li> <li>10. Работа со статистическими источниками в R-studio.</li> <li>11. Методы статистического моделирования в прогнозировании источниками в R-studio.</li> <li>12. Статистический анализ структуры и динамики ВВП России за 2009-20120 годы в R-studio.</li> <li>13. Сравнительный анализ оборота розничной торговли в субъектах Российской Федерации в R-studio.</li> <li>14. Причины недоверия населения к статистическим данным</li> <li>15. Исследовательский проект "Статистика происшествий на дороге" в R-studio.</li> </ol>

#### Задания для тестирования по дисциплине для оценки сформированности компетенций (min 20, max 50 + ссылкой на ЭИОС с тестами)

На какие платформы можно установить язык R?

```
{
-Windows
-MacOS
-Linux
}
```

Как получить справку по функции R?

```
{
-Написать знак вопроса и название функции ?function
-Только через Google или на сайтах вроде stackoverflow
-С помощью функции help
}
```

Одним из важнейших ресурсов администрирования языка R является си-стема

```
{
```

- CRAT
- CRAN (Comprehensive R Archive Network)
- R демократичен

}

Открывая тот или иной файл, вы видите, что этот файл выглядит как набор непонятных символов, такие кракозябры, как это исправить?

{

- изначальный файл пере сохранить, как txt
- установить рабочую директорию
- изменить кодировку на UTF-8

}

Что такое рабочая директория

{

- относительный путь
- папка в операционной системе
- абсолютный путь

}

Как установить путь к рабочей директории?

{

- Выбрать в меню Session, Set working Directory, Choose Directory
- D:\ имя внутренней папки, имя вашей внешней папки
- нажать Ctrl - Enter

}

Что такое Stack Overflow

{

- сообщество форума
- справка по r
- сайт администрирования r

}

Что такое вектор?

{

- это одномерный набор данных, это как книги на полке, у каждого значения вектора есть своя позиция в этой последовательности, и это позиция обозначается квадратными скобками
- числа в ряд
- 100 числовых значений

}

У каждого элемента вектора есть своя позиция в последовательности.

Как эта позиция обозначается в среде R

{

- одним числом
- через оператор вектора
- []

}

Основные способы создания векторов в R

{

- русская «эс» c(1, 6, 3, -11)
- загрузка данных из внешних источников
- my\_vector <- c[1, 0, 3, -10]

}

Векторы с повторяющимися элементами

{

```
[1] 1 1 1 1
```

```
my_vector_2 []
```

```
my_vector_2 <- rep(1, 4)
```

Какая функция в R позволяет повторить вектор или его элемент

```
{  
- rep[,]  
- rep(1:5, 4)  
- Rep(1:5, 4)  
}
```

Выберите правильный вариант задания последовательности с помощью функции seq()

```
{  
- my_vector_5 <- seq(from = 1, to = 10, by = 1)  
- my_vector_6 <- [seq(from=1, to=10, by=0.5)]  
- my_vector_7 <- seq[from=1, to=10, length.out=20]  
}
```

В R многие операции векторизованы, что это значит?

```
{  
- что одно и то же действие очень легко повторить для множества элементов  
- можно получать информацию о свойствах вектора в целом  
- можно прибавлять ко всем элементам одно и то же число  
}
```

Какую функцию можно использовать для проверки типа данных?

```
{  
- функции семейства is.<datatype>:  
- набрав в консоли type  
- с помощью входной переменной  
}
```

Что такое матрица?

```
{  
- матрица, как и вектор, может содержать только элементы одного типа (числовые, строковые, логические и т.д.)  
- два измерения (строки и столбцы)  
- массивы со столбцами и строками  
}
```

Что такое массивы?

```
{  
- это многомерные структуры данных, с количеством измерений 3 и более  
- два измерения (строки и столбцы)  
- фреймы данных  
}
```

Для создания фреймов данных какая используется функция?

```
{  
- data.frame()  
- array()  
- cbind() и rbind()  
}
```

Для сортировки фрейма данных по значению определенного поля, нужно выполнить следующие действия (выбрать из предложенных вариантов)?

```
{  
- необходимо узнать порядок элементов в этом поле df[order(df$lengths), ]  
- необходимо передать условие в первое измерение фрейма df[df$lengths > 40, ]  
- передать название столбцов, которые хранятся как вектор из строк, colnames(df) -  
с("Цвет", "Длина", "Станции", "Плотность")  
}
```

}

Выполняет исходный код на языке высокого уровня «на лету» непосредственно во время работы пр...

::... выполняет исходный код на языке высокого уровня «на лету» непосредственно во время работы пр...::[html]<p>... выполняет исходный код на языке высокого уровня «на лету» непосредственно во время работы программы, без предварительного формирования исполняемого файла.</p>{

- <p>транслятор</p>
- <p>компилятор</p>
- <p>интерпретатор</p>
- <p>парсер</p>

}

Был вызван генератор случайных чисел из библиотеки random: x - random.random(). Проверка (x &g...

::Был вызван генератор случайных чисел из библиотеки random\; x \ - random.random(). Проверка (x &g...::[html]<p>Был вызван генератор случайных чисел из библиотеки random\; x \ - random.random(). Проверка (x &gt; 1)...</p>{

- <p>вернет True</p>
- <p>вернет False</p>
- <p>может вернуть как True, так и False</p>

}

В каких из перечисленных случаях требуется СУБД со свойством расширяемости записей?

::В каких из перечисленных случаев требуется СУБД со свойством расширяемости записей?::[html]<p>В каких из перечисленных случаев требуется СУБД со свойством расширяемости записей?</p>{

- <p>&nbsp;требуется добавлять оценки пользователей музыкальным композициям для целей дальнейшей выдачи рекомендаций</p>
- <p>требуется сохранять все сообщения электронной почты, проходящие через корпоративный узел</p>
- <p>требуется сохранять логи кластера серверов без требования оперативной аналитики</p>
- <p>&nbsp;в проекте требуется индексировать веб-страницы интернета. Каждый месяц аналитики анализируют и добавляют новые признаки, которые вычисляются по проиндексированной веб-странице</p>

}

В какой библиотеке находится структура для табличного представления данных (Data Frame)?

::В какой библиотеке находится структура для табличного представления данных (Data Frame)?::[html]<p>В какой библиотеке находится структура для табличного представления данных (Data Frame)?</p>{

- <p>Pandas</p>
- <p>SciPy</p>
- <p>NumPy</p>
- <p>такой библиотеки нет</p>

}

В каком случае применение Tableau наиболее оправдано

::В каком случае применение Tableau наиболее оправдано::[html]<p>В каком случае применение Tableau наиболее оправдано</p>{

```
-<p>необходимо реализовать гибкое интерактивное визуальное представление
данных</p>
-<p>проведено исследование, результатом которого стала таблица объект-свойства,
необходимо предоставить отчетность</p>
-<p>имеются данные, необходимо более получить ясное понимание этих данных</p>
-<p>не оправдано</p>
}
```

В линейной производственной функции

```
::В линейной производственной функции::[html]<p>В линейной производственной функции</p>{
-<p>установлена фиксированная пропорция ресурсов</p>
-<p>ресурсы свободно заменяемые с постоянной пропорцией</p>
-<p>ресурсы могут заменять друг друга, но цена замены отличается</p>
}
```

В отличие от списков, множества...

```
::В отличие от списков, множества...:[html]<p>В отличие от списков, множества...</p>{
-%50%<p>содержат неповторяющиеся элементы</p>
-<p>строго упорядочены</p>
-%50%<p>содержат элементы в произвольном порядке</p>
-<p>являются неизменяемыми</p>
}
```

В программе выполнен импорт модуля math под именем m. Укажите корректное обращение к функции s...

```
::В программе выполнен импорт модуля math под именем m. Укажите корректное обращение к
функции s...:[html]<p>В программе выполнен импорт модуля math под именем m. Укажите
корректное обращение к функции sin из этого модуля</p>{
-<p>math.sin</p>
-<p>m.sin</p>
-<p>(m)sin</p>
-<p>sin</p>
}
```

В соответствии с рекомендациями Code Style переменные в Python принято называть

```
::В соответствии с рекомендациями Code Style переменные в Python принято
называть::[html]<p>В соответствии с рекомендациями Code Style переменные в Python принято
называть</p>{
-<p>в PascalStyle</p>
-<p>в camelStyle</p>
-<p>в стиле snake_case</p>
-<p>в стиле UPPER_SNAKE_CASE</p>
}
```

В чём состоит свойство расширяемости записей СУБД?

```
::В чём состоит свойство расширяемости записей СУБД?:[html]<p>В чём состоит свойство
расширяемости записей СУБД?</p>{
-<p>СУБД не имеет чёткой структуры, поэтому любую запись можно расширить</p>
-<p>повышение отказоустойчивости системы при добавлении новых записей в
СУБД</p>
-<p>в любую таблицу СУБД можно добавить новую колонку, предварительно изменив
структуру этой таблицы</p>
-<p>СУБД имеет чёткую, но расширяемую структуру, в каждую запись можно добавить
новую колонку, также как и узнать значение любой записи по добавленной колонке</p>
}
```



Величина погрешности характеристики при использовании метода Монте-Карло  
::Величина погрешности характеристики при использовании метода Монте-Карло::[html]<p>Величина погрешности характеристики при использовании метода Монте-Карло</p>{  
-<p>обратно пропорциональна числу испытаний</p>  
-<p>прямо пропорциональна числу испытаний</p>  
-<p>прямо пропорциональна корню из числа испытаний</p>  
-<p>обратно пропорциональна корню из числа испытаний</p>  
}

Выберите корректную запись цикла для перебора значений от 0 до 10 в Python.  
::Выберите корректную запись цикла для перебора значений от 0 до 10 в Python.:[html]<p>Выберите корректную запись цикла для перебора значений от 0 до 10 в Python.</p>{  
-<p>for i \: \ - 0 to 10 do</p>  
-<p>for (int i\ - 0; i &lt;\ - 10; i++)</p>  
-<p>for i in range(11)\:</p>  
-<p>for i in (0, 10)\:</p>  
}

Выберите цикл, который будет бесконечным.  
::Выберите цикл, который будет бесконечным.:[html]<p>Выберите цикл, который будет бесконечным.</p>{  
-<p>for i in range(-1)\:</p>  
-<p>while True\:</p>  
-<p>while False\:</p>  
}

Дистрибутив, содержащий основные библиотеки Python для работы с данными и научных исследований...  
::Дистрибутив, содержащий основные библиотеки Python для работы с данными и научных исследований...:[html]<p>Дистрибутив, содержащий основные библиотеки Python для работы с данными и научных исследований, называется</p>{  
-<p>Anaconda</p>  
-<p>SciPython</p>  
-<p>Jupyter Notebook</p>  
-<p>SciLab</p>  
}

Для каких аргументов функция is.finite вернет >false?  
::Для каких аргументов функция is.finite вернет >false?::[html]<p>Для каких аргументов функция is.finite вернет >false?</p>{  
-%33.33333%<p>&nbsp;(+Inf)</p>  
-%33.33333%<p>&nbsp;NaN</p>  
-%33.33333%<p>&nbsp;NA</p>  
-<p>1</p>  
}

Для каких аргументов функция is.finite вернет >true?  
::Для каких аргументов функция is.finite вернет >true?::[html]<p>Для каких аргументов функция is.finite вернет >true?</p>{  
-<p>NaN</p>  
-%50%<p>&nbsp;1</p>  
-<p>NA</p>  
-%50%<p>(+Inf)</p>  
}

Доверительная вероятность характеризует:

::Доверительная вероятность характеризует\:::[html]<p>Доверительная вероятность характеризует\:</p>{

- <p>степень достоверности полученных оценок</p>
- <p>ширину доверительного интервала</p>
- <p>вероятность попадания случайной величины в доверительный интервал</p>
- <p>вероятность достижения цели моделирования</p>

}

Доверительный интервал случайной величины есть величина...

::Доверительный интервал случайной величины есть величина...:::[html]<p>Доверительный интервал случайной величины есть величина...</p>{

- <p>случайная</p>
- <p>детерминированная</p>
- <p>случайная или детерминированная, в зависимости от объема исходной выборки</p>

}

Известно, что между величинами X и Y существует положительная связь. В каких пределах находится...

::Известно, что между величинами X и Y существует положительная связь. В каких пределах находится...:::[html]<p>Известно, что между величинами X и Y существует положительная связь. В каких пределах находится парный коэффициент корреляции?</p>{

- <p>от -1 до 0</p>
- <p>от 0 до 1</p>
- <p>от -1 до 1</p>

}

Имеет ли Python аналог Data Frame из R

::Имеет ли Python аналог Data Frame из R:::[html]<p>Имеет ли Python аналог Data Frame из R</p>{

- <p>&nbsp;да, библиотека Pandas</p>
- <p>да, библиотека SciPy</p>
- <p>нет</p>
- <p>да, библиотека NumPy</p>

}

Имеется множество  $a = \{1, 3, 5, 7\}$ . После выполнения команды `a.add(7)` множество примет вид...

::Имеется множество  $a = \{1, 3, 5, 7\}$ . После выполнения команды `a.add(7)` множество примет вид...:::[html]<p>Имеется множество  $a = \{1, 3, 5, 7\}$ . После выполнения команды `a.add(7)` множество примет вид...</p>{

- <p> $\{1, 3, 5, 7\}$ </p>
- <p> $\{1, 3, 5\}$ </p>
- <p> $\{1, 3, 5, 7, 7\}$ </p>
- <p>команда вызовет ошибку</p>

}

Имеется множество  $a = \{1, 3, 5, 7\}$ . После выполнения команды `a.remove(9)` множество примет вид...

::Имеется множество  $a = \{1, 3, 5, 7\}$ . После выполнения команды `a.remove(9)` множество примет вид...:::[html]<p>Имеется множество  $a = \{1, 3, 5, 7\}$ . После выполнения команды `a.remove(9)` множество примет вид...</p>{

- <p> $\{1, 3, 5, 7\}$ </p>
- <p> $\{1, 3, 5\}$ </p>
- <p> $\{1, 3, 5, 7, 9\}$ </p>
- <p>команда вызовет ошибку</p>

}

Имеется словарь abc. Цикл for i in abc: перебирает...

```
::Имеется словарь abc. Цикл for i in abc\; перебирает...::[html]<p>Имеется словарь abc. Цикл for i in abc\; перебирает...</p>{  
  -<p>ключи</p>  
  -<p>значения</p>  
  -<p>пары ключ-значение</p>  
  -<p>содержит ошибку</p>  
}
```

Имеется файловая переменная my\_file. Команда my\_file.closed позволяет узнать...

```
::Имеется файловая переменная my_file. Команда my_file.closed позволяет узнать...::[html]<p>Имеется файловая переменная my_file. Команда my_file.closed позволяет узнать...</p>{  
  -<p>имя файла</p>  
  -<p>полный путь к файлу</p>  
  -<p>режим, в котором открыт файл</p>  
  -<p>закрыт файл или открыт</p>  
}
```

Имеется файловая переменная my\_file. Команда my\_file.name позволяет узнать...

```
::Имеется файловая переменная my_file. Команда my_file.name позволяет узнать...::[html]<p>Имеется файловая переменная my_file. Команда my_file.name позволяет узнать...</p>{  
  -<p>имя файла</p>  
  -<p>полный путь к файлу</p>  
  -<p>режим, в котором открыт файл</p>  
  -<p>закрыт файл или открыт</p>  
}
```

Имеется файловая переменная my\_file. Команда my\_file.read() позволяет...

```
::Имеется файловая переменная my_file. Команда my_file.read() позволяет...::[html]<p>Имеется файловая переменная my_file. Команда my_file.read() позволяет...</p>{  
  -<p>открыть файл для чтения</p>  
  -<p>прочитать из файла данные до конца строки</p>  
  -<p>прочитать из файла один байт данных</p>  
  -<p>прочитать весь файл до конца</p>  
}
```

Имена переменных могут включать:

```
::Имена переменных могут включать\:::[html]<p>Имена переменных могут включать\:</p>{  
  -<p>русские буквы</p>  
  -<p>латинские буквы</p>  
  -<p>пробелы</p>  
  -<p>скобки, знаки + \ - ! ? b др.</p>  
  -%25%<p>цифры</p>  
  -%25%<p>знак подчеркивания</p>  
}
```

Инвестиционный фонд интересуется тем, почему часть финансируемых им проектов успешно переходят...

```
::Инвестиционный фонд интересуется тем, почему часть финансируемых им проектов успешно переходят...::[html]<p>Инвестиционный фонд интересуется тем, почему часть финансируемых им проектов успешно переходят на второй год, а часть — нет. К какому типу относится эта задача анализа данных?</p>{  
  -<p>построение решающего правила</p>  
  -<p>классификация</p>  
  -<p>&nbsp;поиск информативных признаков</p>
```

-<p>цензурирование</p>

}

К какому типу шкал относится шкала «очень плохо»-«плохо»-«средне»-«хорошо»-«очень хорошо»?

::К какому типу шкал относится шкала «очень плохо»-«плохо»-«средне»-«хорошо»-«очень хорошо»?::[html]<p>К какому типу шкал относится шкала «очень плохо»-«плохо»-«средне»-«хорошо»-«очень хорошо»?</p>{

-<p>номинальная</p>

-<p>абсолютная</p>

-<p>порядковая</p>

-<p>бинарная</p>

}

К слабоструктурированным относятся задачи, которые содержат

::К слабоструктурированным относятся задачи, которые содержат::[html]<p>К слабоструктурированным относятся задачи, которые содержат</p>{

-<p>как количественные, так и качественные переменные, причём доминируют количественные переменные</p>

-<p>как количественные, так и качественные переменные, причём доминируют качественные переменные</p>

-<p>количественные и качественные переменные в равных долях</p>

}

Как называется метод изучения объекта не непосредственно, а через рассмотрение подобного ему и...

::Как называется метод изучения объекта не непосредственно, а через рассмотрение подобного ему и...::[html]<p>Как называется метод изучения объекта не непосредственно, а через рассмотрение подобного ему и более простого объекта?</p>{

-<p>метод прогнозирования</p>

-<p>метод моделирования</p>

-<p>метод оптимизации</p>

-<p>метод алгоритмизации</p>

-<p>метод деривации</p>

}

Как называется нарушение допущения о независимости остатков?

::Как называется нарушение допущения о независимости остатков?::[html]<p>Как называется нарушение допущения о независимости остатков?</p>{

-<p>Мультиколлинеарность</p>

-<p>Автокорреляция</p>

-<p>Гетероскедастичность</p>

-<p>Гомоскедастичность</p>

}

Как называется нарушение допущения о постоянстве дисперсии остатков?

::Как называется нарушение допущения о постоянстве дисперсии остатков?::[html]<p>Как называется нарушение допущения о постоянстве дисперсии остатков?</p>{

-<p>Мультиколлинеарность</p>

-<p>Автокорреляция</p>

-<p>Гетероскедастичность</p>

-<p>Гомоскедастичность</p>

}

Как правильно присвоить r значение типа данных, имеющих две координаты&nbsp;x - 1&nbsp;и&nbsp;y - ...

```

::Как правильно присвоить p значение типа данных, имеющих две координаты&nbsp;x\ -
1&nbsp;и&nbsp;y\ -...:[html]<p>Как правильно присвоить p значение типа данных, имеющих две
координаты&nbsp;x\ -1&nbsp;и&nbsp;y\ -2?</p>{
  -<p>p(x,y)&lt;-c(1,2)</p>
  -<p>&nbsp;p \ - c(y\ -2, x\ -1)</p>
  -<p>&nbsp;p \ - c(x\ -1, y\ -2)</p>
  -<p>p.x\ -1.y\ -2</p>
}

```

Какая из NoSQL СУБД подходит для организации консистентного и распределённого хранилища?

```

::Какая из NoSQL СУБД подходит для организации консистентного и распределённого
хранилища?::[html]<p>Какая из NoSQL СУБД подходит для организации консистентного и
распределённого хранилища?</p>{

```

- <p>Hbase</p>
- <p>Cassandra</p>
- p>&nbsp;Redis</p>
- p>&nbsp;BigTable</p>

**Практические задачи (min 20, max 50 + ссылку на ЭИОС с электронным изданием, если имеется)**

Раздел дисциплины	Задачи
Современные инструменты и алгоритмы анализа BIG DATA.	<p>ТАА 1</p> <p>Запишите условие проверки неравенства чисел <math>a</math> и <math>b</math> не менее чем <i>три</i> способами.</p> <ol style="list-style-type: none"> <li>1. Напишите программу, которая запрашивает в консоли целое число и определяет, является ли оно чётным или нечетным. Программа должна предварительно определить, является ли введенное число а) числом и б) целым числом.</li> </ol> <p><b>Подсказка:</b> результат конвертации строки в целое число и число с плавающей точкой отличается. Вы можете использовать это для проверки, является ли введенное число целым.</p> <ol style="list-style-type: none"> <li>2. Напишите программу, которая считывает из консоли введенную пользователем строку и выводит в консоль количество символов в этой строке. Вывод оформите следующим образом: "Длина введенной строки равняется ... символам", где вместо многоточия стоит вычисленная длина.</li> <li>3. В программе в виде переменных задайте координаты населенного пункта <math>A(x_1, y_1)</math>, а также дирекционный угол <math>D</math> и расстояние <math>L</math> до населенного пункта <math>B</math>. Напишите код, который определяет координаты населенного пункта <math>B(x_2, y_2)</math>.</li> <li>4. Функция <a href="#">atan2()</a> позволяет найти математический азимут (полярный угол), если известны координаты вектора между двумя точками. Используя эту функцию, напишите программу, которая вычисляет географический азимут между точками <math>A(x_1, y_1)</math> и <math>B(x_2, y_2)</math>. Координаты точек задайте в виде переменных непосредственно в коде.</li> </ol> <p><b>Математический азимут</b> отсчитывается от направления на восток против часовой стрелки. <b>Географический азимут</b> отсчитывается от направления на север по часовой стрелке).</p>
Язык R один из стандартов при работе с данными и	<p>ТАА 2</p> <ol style="list-style-type: none"> <li>1. Создайте вектор temp, в котором хранятся значения среднемесячных температур воздуха в городе Санкт-Петербурге (данные можно взять <a href="#">здесь</a>). Напишите программный код, который вычисляет следующие вектора:</li> </ol>

<p>статистическая обработка данных.</p>	<ul style="list-style-type: none"> <li>▪ <i>количественное изменение</i> температуры от месяца к месяцу (в градусах)</li> <li>▪ <i>качественное изменение</i> температуры от месяца к месяцу ('потепление' или 'похолодание');</li> <li>▪ номера <i>зимних</i> месяцев (со среднемесячной температурой ниже нуля);</li> <li>▪ описательные статистики среднемесячных температур (<i>summary</i>);</li> </ul> <p>Выведите исходные и вычисленные данные в консоль (с пояснением что они означают).</p> <p><b>Подсказка:</b> для вычисления разностей между элементами вектора используйте функцию <code>diff()</code>.</p> <p>2. На местности задан прямоугольник с координатами левого нижнего (<math>x_1, y_1</math>) и правого верхнего (<math>x_2, y_2</math>) угла. Напишите программу, которая размещает внутри этого прямоугольника случайным образом <math>N</math> точек и представляет результат в виде матрицы координат <code>coords</code> с двумя столбцами и <math>N</math> строками. Вызовите в конце программы <code>plot(coords)</code>, чтобы посмотреть на результат. Координаты можно не вводить, а задать прямо в программе в виде переменных.</p> <p><b>Подсказка:</b> координаты случайно размещенных точек имеют равномерное распределение. Вам необходимо сначала сформировать случайные векторы координат <math>X</math> и <math>Y</math>, и после этого объединить их в матрицу.</p> <p>3. Высотная поясность на северном склоне Западного Кавказа, согласно Большой Российской энциклопедии устроена следующим образом:</p> <ul style="list-style-type: none"> <li>▪ до 500 м — степь и лесостепь</li> <li>▪ до 800 м — низкогорные широколиственные леса (дуб, граб)</li> <li>▪ до 1300 м — среднегорные широколиственные леса (бук)</li> <li>▪ до 1600 м — смешанные леса (ель, пихта, бук)</li> <li>▪ до 2300 м — криволесья (береза, бук, клён)</li> <li>▪ до 2500 м — субальпийские и альпийские луга</li> <li>▪ до 3300 м — субнивальная зона (фрагментарная растительность)</li> <li>▪ выше (условно до 5000 м) — гляциально-нивальная зона</li> </ul> <p>Создайте фрейм данных, включающий три столбца: минимальная высота пояса (<math>H_{min}</math>), максимальная высота пояса (<math>H_{max}</math>) и название высотного пояса (<math>Zone</math>). Минимальную высоту надо вычислить на основе максимальной, приняв, что для нижнего пояса она условно равна 400 м.</p> <p>Напишите программу, которая просит пользователя ввести высоту и возвращает высотный пояс, соответствующую введенной высоте (достаточно вывести строчку фрейма данных).</p> <p><b>Подсказка:</b> Организуйте обход строчек фрейма данных с помощью цикла от 11 до NN, где NN — количество строк. Искомый пояс будет найден, как только введенное значение станет меньше чем <math>H_{max}</math>. После этого можно вывести результат на экран. Если введенное значение больше максимума в столбце <math>H_{max}</math> или меньше 400, программа должна выдавать ошибку.</p> <p>4. <b>[advanced]</b> Решите задачу №3, используя только операции над векторами и поиск элементов, и не используя циклы.</p> <p>5. <b>[advanced]</b> Модифицируйте программу, написанную для решения задачи №2 таким образом, чтобы запретить точкам сближаться более чем на заданное расстояние <math>k</math> (это называется <i>регулярным распределением</i> с расстоянием ингибции <math>k</math>). Сохраните результат в</p>
---	---

	<p>виде <i>фрейма данных</i> points со столбцами <math>X</math>, <math>Y</math> и <math>D</math>, где <math>D</math> – это расстояние до ближайшей точки. Выведите верхние строчки полученной таблицы в консоль, а также полученные точки с помощью команды <code>plot(coords\$X, coords\$Y)</code>.</p> <p><b>Подсказка:</b> вам придется генерировать в цикле по одной точке и проверять условие на каждой итерации до тех пор, пока вы не наберете требуемое количество точек. Задавайте значение <math>k</math> малым по отношению к размерам прямоугольника, чтобы избежать излишне долгого выполнения программы.</p> <p>6. Постройте для набора данных quakes пакета datasets гистограммы распределения глубин и магнитуд, а также диаграмму рассеяния для двух этих характеристик.</p> <p>7. На портале открытых данных Самарской области есть данные о распределении площади лесов и запасов древесины по преобладающим породам и группам возраста. Скачайте эти данные в виде таблицы CSV и постройте по ним круговую и столбчатую диаграмму для категории Площадь земель, занятых лесными насаждениями (покрытых лесной растительностью), всего. Подберите цвета, попробуйте изменить ориентировку столбцов на горизонтальную, а для круговой диаграммы поменять угол поворота.</p> <p>8. Используя данные<sup>3</sup> по балансу масс ледника Гарабаши, постройте график с тремя кривыми (аккумуляции, абляции и кумулятивного баланса) за период 1981 по 2017 г. Добавьте на график легенду. Обратите внимание на то, что таблица содержит агрегирующие строки (1982-1997, 1998-2017, 1982-2017), которые вам необходимо предварительно исключить.</p> <p>Подсказка: Чтобы построить кривую кумулятивного баланса, используйте функцию <code>cumsum</code>.</p> <p>9. Таблица storms из пакета dplyr содержит данные трекинга тропических циклонов с 1975 по 2015 год. Извлеките из нее данные по одному выбранному циклону и постройте на их основе график, показывающий трек прохождения циклона в системе координат давление (<math>X</math>) — скорость (<math>Y</math>). Каждое наблюдение циклона визуализируйте в виде кружка, цвет которого зависит от класса циклона (переменная status), а размер — от диаметра территории, на которой наблюдаются скорости ветра класса тропического шторма и сильнее. Нанесите поверх кружков значение диаметра текстовой подписью. Соедините кружки линиями со стрелкой посередине, чтобы была понятна история жизни циклона.</p>
--	--

## 6.5. Оценочные материалы для промежуточной аттестации

### Фонд вопросов для проведения промежуточного контроля в форме экзамена

Раздел дисциплины	Вопросы
Современные инструменты и алгоритмы анализа BIG DATA.	<ol style="list-style-type: none"> <li>1. В чем принципиальное отличие концепции Big Data от традиционного подхода BI?</li> <li>2. Понятие явной (выраженной) и скрытой (структурной) информации.</li> <li>3. Определение контент-анализа.</li> <li>4. Каковы основные понятия контент-анализа?</li> <li>5. Какие существуют виды контент-анализа?</li> <li>6. Какие существуют этапы контент-анализа?</li> <li>7. Каковы основные признаки, характеризующие «Большие данные»?</li> <li>8. Модели развертывания облачных хранилищ.</li> </ol>

	<p>9. Модели обслуживания облачных хранилищ.  10. Постановка и описание проблемы «последней мили».  11. Безопасность, производительность и надежность при работе с облачными данными.  12. Экономическая составляющая облачных подходов.  13. Способы машинного обучения.  14. Основные фазы обработки «больших данных».</p>
<p>Язык R один из стандартов при работе с данными и статистическая обработка данных.</p>	<p>15. Работа с комплексными числами в программе R (Rstudio). Решение систем линейных уравнений с комплексными коэффициентами с использованием matrix() и solve().  16. Параметры выборки: медиана, мода, среднее арифметическое. Примеры расчёта.  17. Линейная регрессия в R  18. Типы данных в R  19. Основные операторы  20. Вектора, матрицы, дата фреймы  21. Функции  22. Функции. Основные распределения  23. Работа с внешними пакетами  24. Загрузка встроенных в пакеты датасетов  25. Импорт данных из файла формата csv  26. Импорт данных из файла формата MS Excel  27. Оценивание линейной регрессии  28. Прогнозирование по оценённой модели линейной регрессии  29. Тестирование гипотез о коэффициент  30. F-тест для коэффициентов регрессии  31. ДИАГНОСТИЧЕСКИЕ ТЕСТЫ</p>

#### 6.6. Шкалы и критерии оценивания по формам текущего контроля и промежуточной аттестации

##### Шкала и критерии оценивания

Оценка	Критерии оценивания для мероприятий контроля с применением 4-х балльной системы
«отлично»	Повышенный УК-1.1, УК-1.2, УК-1.3, ПК-4.1, ПК-4.2, ПК-4.3, ПК-3.1, ПК-3.2, ПК-3.3
«хорошо»	Стандартный УК-1.1, УК-1.2, УК-1.3, ПК-4.1, ПК-4.2, ПК-4.3, ПК-3.1, ПК-3.2, ПК-3.3
«удовлетворительно»	Пороговый УК-1.1, УК-1.2, УК-1.3, ПК-4.1, ПК-4.2, ПК-4.3, ПК-3.1, ПК-3.2, ПК-3.3
«неудовлетворительно»	Результаты обучения не сформированы на пороговом уровне