

Документ подписан электронной подписью
Информация о владельце:
ФИО: Хоружий Людмила Ивановна
Должность: Директор института экономики и управления АПК
Дата подписания: 15.07.2022 19:17:36
Уникальный программный ключ:
1e90b132d9b04dce67585160b015ddd12cb1e6a9



МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
**«РОССИЙСКИЙ ГОСУДАРСТВЕННЫЙ АГРАРНЫЙ УНИВЕРСИТЕТ –
МСХА имени К.А. ТИМИРЯЗЕВА»**
(ФГБОУ ВО РГАУ - МСХА имени К.А. Тимирязева)

Институт экономики и управления АПК
Кафедра статистики и кибернетики

УТВЕРЖДАЮ:
Директор института экономики и управ-
ления АПК
Л.И. Хоружий
“ 15 ” 2022 г.

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ
ФТД.02 Обработка данных в среде R**

для подготовки бакалавров

ФГОС ВО

Направление: 09.03.02 «Информационные системы и технологии»

Направленность:

Большие данные и машинное обучение (Machine Learning & Big Data)

Курс 3,4

Семестр 6,7

Форма обучения очная

Год начала подготовки 2022

Москва, 2022

Разработчик (и): Харитонов А.Е., к.э.н., доцент
(ФИО, ученая степень, ученое звание)

«26» 08 2022г.

Рецензент: Коломеева Е.С., к.э.н.
(ФИО, ученая степень, ученое звание)

«26» 08 2022г.

Программа составлена в соответствии с требованиями ФГОС ВО, профессионального стандарта и учебного плана по направлению подготовки 09.03.02 «Информационные системы и технологии».

Программа обсуждена на заседании кафедры статистики и кибернетики протокол № 11 от «26» августа 2022 г.

И.о.зав. кафедрой Уколова А.В., к.э.н., доцент
(ФИО, ученая степень, ученое звание)

«26» 08 2022г.

Согласовано:

Председатель учебно-методической комиссии института экономики и управления АПК
Корольков А.Ф., к.э.н., доцент
(ФИО, ученая степень, ученое звание)

«26» 08 2022г.

И.о.заведующего выпускающей кафедрой
статистики и кибернетики
Уколова А.В., к.э.н., доцент
(ФИО, ученая степень, ученое звание)

«26» 08 2022г.

/Заведующий отделом комплектования ЦНБ

Ермилова Э.В.
(подпись)

СОДЕРЖАНИЕ

АННОТАЦИЯ	4
1. ЦЕЛЬ ОСВОЕНИЯ ДИСЦИПЛИНЫ	6
2. МЕСТО ДИСЦИПЛИНЫ В УЧЕБНОМ ПРОЦЕССЕ	6
3. ПЕРЕЧЕНЬ ПЛАНИРУЕМЫХ РЕЗУЛЬТАТОВ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ, СООТНЕСЕННЫХ С ПЛАНИРУЕМЫМИ РЕЗУЛЬТАТАМИ ОСВОЕНИЯ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ	6
4. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ	6
4.1 РАСПРЕДЕЛЕНИЕ ТРУДОЁМКОСТИ ДИСЦИПЛИНЫ ПО ВИДАМ РАБОТ	7
ПО СЕМЕСТРАМ	7
4.2 СОДЕРЖАНИЕ ДИСЦИПЛИНЫ.....	10
4.3 ПРАКТИЧЕСКИЕ ЗАДАНИЯ.....	12
5. ОБРАЗОВАТЕЛЬНЫЕ ТЕХНОЛОГИИ	14
6. ТЕКУЩИЙ КОНТРОЛЬ УСПЕВАЕМОСТИ И ПРОМЕЖУТОЧНАЯ АТТЕСТАЦИЯ ПО ИТОГАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ	14
6.1. ТИПОВЫЕ КОНТРОЛЬНЫЕ ЗАДАНИЯ ИЛИ ИНЫЕ МАТЕРИАЛЫ, НЕОБХОДИМЫЕ ДЛЯ ОЦЕНКИ ЗНАНИЙ, УМЕНИЙ И НАВЫКОВ И (ИЛИ) ОПЫТА ДЕЯТЕЛЬНОСТИ	14
6.2. ОПИСАНИЕ ПОКАЗАТЕЛЕЙ И КРИТЕРИЕВ КОНТРОЛЯ УСПЕВАЕМОСТИ, ОПИСАНИЕ ШКАЛ ОЦЕНИВАНИЯ	17
7. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ	18
7.1 ОСНОВНАЯ ЛИТЕРАТУРА	18
7.2 ДОПОЛНИТЕЛЬНАЯ ЛИТЕРАТУРА.....	18
7.3 МЕТОДИЧЕСКИЕ УКАЗАНИЯ, РЕКОМЕНДАЦИИ И ДРУГИЕ МАТЕРИАЛЫ К ЗАНЯТИЯМ.....	19
8. ПЕРЕЧЕНЬ РЕСУРСОВ ИНФОРМАЦИОННО-ТЕЛЕКОММУНИКАЦИОННОЙ СЕТИ «ИНТЕРНЕТ», НЕОБХОДИМЫХ ДЛЯ ОСВОЕНИЯ ДИСЦИПЛИНЫ	19
9. ПЕРЕЧЕНЬ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ И ИНФОРМАЦИОННЫХ СПРАВОЧНЫХ СИСТЕМ	19
10. ОПИСАНИЕ МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЙ БАЗЫ, НЕОБХОДИМОЙ ДЛЯ ОСУЩЕСТВЛЕНИЯ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА ПО ДИСЦИПЛИНЕ	20
11. МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ ОБУЧАЮЩИМСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ	20
12. МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ ПРЕПОДАВАТЕЛЯМ ПО ОРГАНИЗАЦИИ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ	21

Аннотация

рабочей программы учебной дисциплины

ФТД.02 «Обработка данных в среде R»

для подготовки бакалавров по направлению 09.03.02 «Информационные системы и технологии» по направленности Большие данные и машинное обучение (Machine Learning & Big Data) и

Цель освоения дисциплины: Целью дисциплины «Обработка данных в среде R» является освоение студентами теоретических и практических знаний и приобретение умений и навыков в области использования современных информационных технологий для осуществления анализ и синтеза информации для решения поставленных задач профессиональной деятельности.

Место дисциплины в учебном плане: дисциплина включена в факультативные дисциплины учебного плана по направлению подготовки 09.03.02 «Информационные системы и технологии».

Требования к результатам освоения дисциплины: в результате освоения дисциплины формируются следующие компетенции (индикаторы): УК-1 (УК-1.3); ПКос-8 (ПКос-8.2; ПКос-8.3).

Краткое содержание дисциплины:

Установка и настройка программы. Установка RStudio. Общие принципы работы. Подключение пакетов. Console. Script. Environment. Помощь в работе. Справка. Рабочее пространство. Корневой каталог. Набор данных. Структуры данных. Векторы. Создание. Изменение. Обращение к элементам. Матрицы. Способы создания. Операции с матрицами. Таблицы данных. Attach, detach и with. Названия строк. Массивы. Факторы. Списки. Ввод данных. Ввод данных с клавиатуры. Импорт данных из текстового файла с разделителями. Импорт данных из XML-файлов. Извлечение данных из веб-страниц. Импорт данных из Stata Импорт данных из систем управления базами данных. Интерфейс ODBC. Импорт данных при помощи Stat/Transfer. Создание диаграммы. Графические параметры. Символы и линии. Цвета. Характеристики текста. Размеры диаграммы и полей. Добавление текста, настройка параметров осей и условных обозначений. Заголовки. Легенда. Аннотации. Подписи с математическими символами. Объединение диаграмм. Полный контроль над расположением диаграмм. Работа с датами и пропущенными значениями. Преобразование типов данных в другой. Создание и перекодирование переменных. Переименование переменных. Сортировка, объединение и разделение наборов данных. Выбор и исключение переменных из анализа. Преобразование дат в текстовые переменные. Сортировка данных. Объединение наборов данных. Разделение наборов данных на составляющие. Случайные выборки. Использование команд SQL для преобразования таблиц данных. Математические и статистические функции. Функции распределения. Генерация случайных чисел. Текстовые функции. Циклы и исполнение команд при условии. Пользовательские функции. Применение функций к матрицам и таблицам данных. Управление выполнением команд. Способы объединять и преобразовывать данные. Операторы цикла и условия. Создание собственных функций в R. Операторы цикла и условия в R .

Оператор if . Оператор ifelse . Оператор for . Оператор while . Операторы repeat, break и next . Оператор switch . Стандартная форма задания функции в R. Аргумент. Формальные аргументы, локальные переменные и свободные переменные . Полная форма задания функции в R. Сильное присваивание в. Команды apply(), sapply() и lapply() . Примеры написания функций в R с использованием управляющих конструкций. Описательные статистики. Таблицы частот и таблицы сопряженности. Корреляция и ковариация. Тесты Стьюдента. Непараметрические методы. Визуализация групповых различий. Создание и интерпретация линейных моделей. Оценка адекватности допущений, сделанных при построении модели. Выбор между альтернативными моделями. Обнаружение пропущенных данных. Классификация типов пропущенных данных. Исследование структуры пропущенных данных Визуализация закономерностей в пропущенных данных. Анализ полных наблюдений. Множественное восстановление пропущенных данных.

Общая трудоемкость дисциплины составляет: 4 зачетные единицы (144 часа).

Промежуточный контроль: зачет в 6 и 7 семестрах.

1. Цель освоения дисциплины

Целью дисциплины «Обработка данных в среде R» является освоение студентами теоретических и практических знаний и приобретение умений и навыков в области использования современных информационных технологий для осуществления анализ и синтеза информации для решения поставленных задач профессиональной деятельности.

2. Место дисциплины в учебном процессе

Дисциплина «Обработка данных в среде R» включена в факультативные дисциплины учебного плана. Дисциплина «Обработка данных в среде R» реализуется в соответствии с требованиями ФГОС ВО, профессионального стандарта, ОПОП ВО и Учебного плана по направлению 09.03.02 «Информационные системы и технологии».

Предшествующими курсами, на которых непосредственно базируется дисциплина «Обработка данных в среде R» являются «Линейная алгебра», «Математический анализ», «Математическая статистика», «Теория вероятностей», «Алгоритмизация и программирование», «Технологии хранения и управления данными».

Дисциплина «Обработка данных в среде R» является основополагающей для изучения следующих дисциплин: «Построение, обучение и оптимизация моделей машинного обучения», «Построение, обучение и оптимизация моделей машинного обучения на иностранном языке», «Разработка средств интеграции и поддержки готового решения для анализа больших данных», «Разработка средств интеграции и поддержки готового решения для анализа больших данных на иностранном языке».

Рабочая программа дисциплины «Обработка данных в среде R» для инвалидов и лиц с ограниченными возможностями здоровья разрабатывается индивидуально с учетом особенностей психофизического развития, индивидуальных возможностей и состояния здоровья таких обучающихся.

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

Образовательные результаты освоения дисциплины обучающимся, представлены в таблице 1.

4. Структура и содержание дисциплины

4.1 Распределение трудоёмкости дисциплины по видам работ по семестрам

Общая трудоёмкость дисциплины составляет 4 зач.ед. (144 часа), их распределение по видам работ семестрам представлено в таблице 2.

Таблица 1

Требования к результатам освоения учебной дисциплины

№ п/п	Код компетенции	Содержание компетенции (или её части)	Индикаторы компетенций	В результате изучения учебной дисциплины обучающиеся должны:		
				знать	уметь	владеть
1.	УК-1	Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач	УК-1.3 Владеть: методами поиска, сбора и обработки, критического анализа и синтеза информации; методикой системного подхода для решения поставленных задач			методами сбора, обработки, анализа и синтеза информации на языке программирования R
2.	ПКос-8	Способность проводить анализ данных с использованием информационных технологий в области сельского хозяйства, экономики, бухгалтерского учета, статистики, финансов и др.	ПКос-8.2 Уметь: собирать информацию для проведения анализа; устанавливать причинно-следственные связи между признаками; выбирать и применять, в том числе с использованием современных информационных технологий, методы анализа данных; делать выводы на основе проведенного анализа данных		применять язык программирования R для сбора и обработки данных, устанавливать причинно-следственные связи между признаками, выбирать и применять методы анализа данных, делать выводы на основе проведенного анализа данных	
3.			ПКос-8.3 Владеть: методологией и на-			навыками проведения анализа данных в среде

			выками проведения анализа данных с использованием информационных технологий в области экономики, в том числе экономики сельского хозяйства			Р в области экономики, в том числе экономики сельского хозяйства
--	--	--	--	--	--	--

Таблица 2

Распределение трудоёмкости дисциплины по видам работ по семестрам

Вид учебной работы	Трудоёмкость		
	час. всего/*	в т.ч. по семестрам	
		№ 6	№ 7
Общая трудоёмкость дисциплины по учебному плану	144	72	72
1. Контактная работа:	32,5/8	16,25/4	16,25/4
Аудиторная работа	32,5/8	16,25/4	16,25/4
<i>в том числе:</i>			
<i>практические занятия (ПЗ)</i>	32/8	16/4	16/4
<i>контактная работа на промежуточном контроле (КРА)</i>	0,5	0,25	0,25
2. Самостоятельная работа (СРС)	111,5	55,75	55,75
<i>самостоятельное изучение разделов, самоподготовка (проработка и повторение лекционного материала и материала учебников и учебных пособий, подготовка к практическим занятиям и т.д.)</i>	93,5	46,75	46,75
<i>Подготовка к зачету (контроль)</i>	18	9	9
Вид промежуточного контроля:		зачёт	зачёт

* в том числе практическая подготовка.

4.2 Содержание дисциплины

Таблица 3

Тематический план учебной дисциплины

Наименование разделов и тем дисциплин (укрупнёно)	Всего	Аудиторная работа		Внеаудиторная работа СР
		ПЗ всего /*	ПКР	
Тема 1. Введение в R	12	2		10
Тема 2. Создание набора данных	10/2	4/2		6
Тема 3. Ввод и импорт данных	6	2		4
Тема 4. Графика в R	8	2		6
Тема 5. Основы управления данными	14	2		12
Тема 6. Числовые и текстовые функции	21,75/2	4/2		17,75
Контактная работа на промежуточном контроле (КРА)	0,25		0,25	
Всего за 6 семестр	72/4	16	0,25	55,75
Тема 7. Создание собственных функций	40/2	8/2		32
Тема 8. Основные методы статистической обработки данных.	16/2	4/2		12
Тема 9. Работа с пропущенными данными	15,75	4		11,75
Контактная работа на промежуточном контроле (КРА)	0,25		0,25	
Всего за 7 семестр	72/4	16	0,25	55,75
Итого по дисциплине	144/8	32	0,5	111,5

* в том числе практическая подготовка

Тема 1. Введение в R

Установка и настройка программы. Установка RStudio. Общие принципы работы. Подключение пакетов. Console. Script. Environment. Помощь в работе. Справка. Рабочее пространство. Корневой каталог.

Тема 2. Создание набора данных

Набор данных. Структуры данных. Векторы. Создание. Изменение. Обращение к элементам. Матрицы. Способы создания. Операции с матрицами. Таблицы данных. Attach, detach и with. Названия строк. Массивы. Факторы. Списки.

Тема 3. Ввод и импорт данных

Ввод данных. Ввод данных с клавиатуры. Импорт данных из текстового файла с разделителями. Импорт данных из XML-файлов. Извлечение данных из веб-страниц. Импорт данных из Stata Импорт данных из систем управления базами данных. Интерфейс ODBC. Импорт данных при помощи Stat/Transfer.

Тема 4. Графика в R

Создание диаграммы. Графические параметры. Символы и линии. Цвета. Характеристики текста. Размеры диаграммы и полей. Добавление текста, настройка параметров осей и условных обозначений. Заголовки. Легенда. Аннотации. Подписи с математическими символами. Объединение диаграмм. Полный контроль над расположением диаграмм.

Тема 5. Основы управления данными

Работа с датами и пропущенными значениями. Преобразование типов данных в другой. Создание и перекодирование переменных. Переименование переменных. Сортировка, объединение и разделение наборов данных. Выбор и исключение переменных из анализа. Преобразование дат в текстовые переменные. Сортировка данных. Объединение наборов данных. Разделение наборов данных на составляющие. Случайные выборки. Использование команд SQL для преобразования таблиц данных.

Тема 6. Числовые и текстовые функции

Математические и статистические функции. Функции распределения. Генерация случайных чисел. Текстовые функции. Циклы и исполнение команд при условии. Пользовательские функции. Применение функций к матрицам и таблицам данных. Управление выполнением команд. Способы объединять и преобразовывать данные.

Тема 7. Создание собственных функций

Операторы цикла и условия. Создание собственных функций в R. Операторы цикла и условия в R . Оператор if . Оператор ifelse . Оператор for . Оператор while . Операторы repeat, break и next . Оператор switch . Стандартная форма задания функции в R. Аргумент. Формальные аргументы, локальные переменные и свободные переменные . Полная форма задания функции в R. Сильное присваивание в. Команды apply(), sapply() и lapply() . Примеры написания функций в R с использованием управляющих конструкций.

Тема 8. Основные методы статистической обработки данных.

Описательные статистики. Таблицы частот и таблицы сопряженности. Корреляция и ковариация. Тесты Стьюдента. Непараметрические методы. Визуализация групповых различий. Создание и интерпретация линейных моделей. Оцен-

ка адекватности допущений, сделанных при построении модели. Выбор между альтернативными моделями.

Тема 9. Работа с пропущенными данными

Обнаружение пропущенных данных. Классификация типов пропущенных данных. Исследование структуры пропущенных данных. Визуализация закономерностей в пропущенных данных. Анализ полных наблюдений. Множественное восстановление пропущенных данных.

4.3 Практические занятия

Таблица 4

Содержание практических занятий и контрольные мероприятия

№ п/п	№ раздела	№ и название практических занятий	Формируемые компетенции (индикаторы)	Вид контрольного мероприятия	Кол-во Часов/ из них практическая подготовка
1	Тема 1. Введение в R	Практическая работа № 1. Интерфейс и настройка R и RStudio	ПКос-8.2 ПКос-8.3	Защита работы	2
2	Тема 2. Создание набора данных	Практическая работа № 2. Вектора и матрицы	ПКос-8.2 ПКос-8.3	Защита работы	2/2
		Практическая работа № 3. Массивы. Факторы. Списки.	ПКос-8.2 ПКос-8.3	Защита работы	2
3	Тема 3. Ввод и импорт данных	Практическая работа № 4. Методы ввода и импорта данных	УК-1.3 ПКос-8.2 ПКос-8.3	Защита работы	2
4	Тема 4. Графика в R	Практическая работа № 5. Построение графиков	УК-1.3 ПКос-8.2 ПКос-8.3	Защита работы	2
5	Тема 5. Основы управления данными	Практическая работа № 6. Сортировка, объединение и разделение наборов данных	УК-1.3 ПКос-8.2 ПКос-8.3	Защита работы	2
6	Тема 6. Числовые и текстовые функции	Практическая работа № 7. Применение математических и статистических функции	УК-1.3 ПКос-8.2 ПКос-8.3	Защита работы	2/2
		Контрольная работа № 1	УК-1.3 ПКос-8.2 ПКос-8.3	Письменная работа	2
7	Тема 7. Создание собственных функций	Практическая работа № 8. Основные операторы и аргументы	ПКос-8.2 ПКос-8.3	Защита работы	2
		Практическая работа № 9. Функции с ограниченным числом переменных	ПКос-8.2 ПКос-8.3	Защита работы	2
		Практическая работа № 10. Функции с неограниченным числом переменных	ПКос-8.2 ПКос-8.3	Защита работы	2

№ п/п	№ раздела	№ и название практических занятий	Формируемые компетенции (индикаторы)	Вид контрольного мероприятия	Кол-во Часов/ из них практическая подготовка
		Практическая работа № 11. Создание функций для определенного класса	ПКос-8.2 ПКос-8.3	Защита работы	2/2
8	Тема 8. Основные методы статистической обработки данных.	Практическая работа № 12 Создание функций для расчета описательных статистик	ПКос-8.2 ПКос-8.3	Защита работы	2/2
		Практическая работа № 13 Создание функций для применения статистических гипотез	УК-1.3 ПКос-8.2 ПКос-8.3	Защита работы	2
9	Тема 9. Работа с пропущенными данными	Практическая работа № 14 Методы обработки пропущенных данных	УК-1.3 ПКос-8.2 ПКос-8.3	Защита работы	2
		Контрольная работа № 2	УК-1.3 ПКос-8.2 ПКос-8.3	Письменная работа	2

Таблица 5

Перечень вопросов для самостоятельного изучения дисциплины

№ п/п	№ раздела и темы	Перечень рассматриваемых вопросов для самостоятельного изучения
1.	Тема 1. Введение в R	Справка. Рабочее пространство. Корневой каталог. (ПКос-8.2, ПКос-8.3)
2.	Тема 2. Создание набора данных	Таблицы данных. Attach, detach и with. (ПКос-8.2, ПКос-8.3)
3	Тема 3. Ввод и импорт данных	Импорт данных из систем управления базами данных. Интерфейс ODBC. Импорт данных при помощи Stat/Transfer. (УК-1.3, ПКос-8.2, ПКос-8.3)
4	Тема 4. Графика в R	Заголовки. Легенда. Аннотации. Подписи с математическими символами. Объединение диаграмм. Полный контроль над расположением диаграмм. (УК-1.3, ПКос-8.2, ПКос-8.3)
5	Тема 5. Основы управления данными	Объединение наборов данных. Разделение наборов данных на составляющие. Случайные выборки. Использование команд SQL для преобразования таблиц данных. (УК-1.3, ПКос-8.2, ПКос-8.3)
6	Тема 6. Числовые и текстовые функции	Применение функций к матрицам и таблицам данных. Управление выполнением команд. Способы объединять и преобразовывать данные. (УК-1.3, ПКос-8.2, ПКос-8.3)
9	Тема 7. Создание собственных функций	Сильное присваивание в. Команды apply(), sapply() и lapply(). Примеры написания функций в R с использованием управляющих конструкций. (ПКос-8.2, ПКос-8.3)
10	Тема 8. Основные методы статистической обработки данных.	Оценка адекватности допущений, сделанных при построении модели. Выбор между альтернативными моделями. (УК-1.3, ПКос-8.2, ПКос-8.3)

№ п/п	№ раздела и темы	Перечень рассматриваемых вопросов для самостоятельного изучения
11	Тема 9. Работа с пропущенными данными	Анализ полных наблюдений. Множественное восстановление пропущенных данных. (УК-1.3, ПКос-8.2, ПКос-8.3)

5. Образовательные технологии

Таблица 6

Применение активных и интерактивных образовательных технологий

№ п/п	Тема и форма занятия	Наименование используемых активных и интерактивных образовательных технологий
1.	Практическая работа № 5. Построение графиков	ПЗ Деловая игра

6. Текущий контроль успеваемости и промежуточная аттестация по итогам освоения дисциплины

6.1. Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений и навыков и (или) опыта деятельности

6.1. Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений и навыков и (или) опыта деятельности

Контрольная работа №1

Вариант 1

1. Создать вектор v (2,5,8,2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2)

Создать вектор b из 20 случайных чисел в диапазоне от 1 до 100 без повторений.

Найти: для вектора b : среднюю (по формуле и по встроенной функции), среднеквадратическое отклонение (округлить в большую сторону), минимальное и максимальное значение, количество элементов.

Вектор v отсортировать в порядке возрастания. Найти значения элементов по ранжированному ряду с номерами 2 и 8.

Рассчитать ковариацию между векторами v и b ($\text{cov} = \overline{xy} - \bar{x} * \bar{y}$).

1. Создайте матрицу A :

2 3
3 5
4 8

Добавьте к этой матрице столбец справа с номером Вашего варианта.

Создайте матрицу B :

6 6 2
4 7 6
5 8 7

Найти: $|A * B + A^T * B^{-1}|$

Практическая работа № 2. Вектора и матрицы

Загрузите RStudio. Создайте вектор, состоящий из 100 случайных значений в соответствии с нормальным законом распределения. По вектору рассчитать показатели центральной тенденции и вариации. Выбрать 7 и 12 элементы вектора. Создать второй вектор, состоящих из 100 значений в диапазоне от 1 до 1000 с шагом 10. Найти сумму произведений элементов векторов. Написать ответ о проделанной работе с выводами.

Создайте 2 матрицы (3 на 3), состоящую из случайных значений. Найдите сумму матриц, произведение, разность матриц, обратные, транспонированные матрицы. Написать ответ о проделанной работе с выводами.

Вопросы к защите:

1. Векторы. Создание. Изменение. Обращение к элементам.
2. Матрицы. Способы создания. Операции с матрицами.

Практическая работа № 3. Массивы. Факторы. Списки.

Загрузите RStudio. Создайте 2 случайных вектора. Преобразуйте из в массив, фактор и в список. Объедините два вектора в таблицу. Присвоить каждому столбцу таблицы название. Обратитесь к 5 элементу 2-го столбца. По показателям таблицы найти 3-ий столбец, как сумму предыдущих. Написать ответ о проделанной работе с выводами.

Вопросы к защите:

1. Массивы. Создание и обращения к элементам.
2. Факторы. Сущность. Использование.
3. Списки. Обращение к элементам.

Практическая работа № 4 . Методы ввода и импорта данных

Загрузите RStudio. Создайте набор исходных показателей в формате .txt. Загрузите эту таблицу используя импорт данных и console. Откройте загруженные данные. Представьте исходные показатели в формате .xls. Загрузите данные в RStudio. Написать ответ о проделанной работе с выводами.

Вопросы к защите:

1. Ввод данных.
2. Ввод данных с клавиатуры.
3. Импорт данных из текстового файла с разделителями.
4. Таблицы данных. Attach, detach и with. Названия строк.

Практическая работа № 5. Построение графиков.

Загрузите RStudio. Загрузите исходные показатели. Требуется построить огиву, полигон, гистограмму, круговую диаграмму, диаграмму рассеяния и лица Чернова. Написать ответ с выводами по каждому графику.

Вопросы к защите:

1. Графика в R
2. Функция plot
3. Гистограммы.

4. Круговые диаграммы.
5. Текстовые функции.

Вопросы к зачету (6 семестр)

1. Набор данных.
2. Структуры данных.
3. Векторы. Создание. Изменение. Обращение к элементам.
4. Матрицы. Способы создания. Операции с матрицами.
5. Таблицы данных. Attach, detach и with. Названия строк.
6. Массивы. Факторы. Списки.
7. Ввод данных. Ввод данных с клавиатуры.
8. Импорт данных из текстового файла с разделителями.
9. Графика в R
10. Преобразование типов данных.
11. Создание и перекодирование переменных.
12. Сортировка, объединение и разделение наборов данных.
13. Математические и статистические функции.
14. Функции распределения. Генерация случайных чисел.
15. Текстовые функции.

Вопросы к зачету (7 семестр)

1. Операторы цикла и условия.
2. Операторы цикла и условия в R .
3. Оператор if . Оператор ifelse . Оператор for .
4. Оператор while. Операторы repeat, break и next .
5. Оператор switch .
6. Команды apply(), sapply() и lapply() .
7. Примеры написания функций в R с использованием управляющих конструкций.
8. Описательные статистики.
9. Таблицы частот и таблицы сопряженности.
10. Непараметрические методы.
11. Визуализация групповых различий.
12. Обнаружение пропущенных данных.
13. Классификация типов пропущенных данных.

6.2. Описание показателей и критериев контроля успеваемости, описание шкал оценивания

Текущий контроль знаний, умений и навыков проводится в форме контрольных работ с расчетными задачами, а также предусмотрена защита индивидуальных задач, выполняемых каждым студентом на практических занятиях. Ликвидация студентами текущих задолженностей производится также в форме

выполнения индивидуальной задачи по соответствующей теме и дальнейшей ее защиты преподавателю кафедры.

Итоговая оценка учитывает результаты рейтинговой системы контроля знаний (вклад 50%), результаты контрольных работ по всем темам (вклад 50%)
Критерии выставления оценок по системе:

- 0-59 % от максимального количества баллов – «незачтено»;
- 60 и более %– «зачтено».

7. Учебно-методическое и информационное обеспечение дисциплины

7.1 Основная литература

1. Гришин, В. А. Основы программирования на языке R : учебно-методическое пособие / В. А. Гришин. — Нижний Новгород : ННГУ им. Н. И. Лобачевского, 2021. — 67 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/191498> (дата обращения: 15.11.2022). — Режим доступа: для авториз. пользователей.
2. Митина, О. А. Языки программирования для статистической обработки данных (R) : учебное пособие / О. А. Митина. — Москва : РТУ МИРЭА, 2020. — 191 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/163912> (дата обращения: 15.11.2022). — Режим доступа: для авториз. пользователей.
3. Гришин, В. А. Методы обработки данных и моделирование на языке R : учебно-методическое пособие / В. А. Гришин, М. С. Тихов. — Нижний Новгород : ННГУ им. Н. И. Лобачевского, 2019. — 54 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/144653> (дата обращения: 15.11.2022). — Режим доступа: для авториз. пользователей.
4. Лукьянов, П. Б. Графический анализ на языке R : учебно-методическое пособие / П. Б. Лукьянов. — Москва : Прометей, 2022. — 48 с. — ISBN 978-5-00172-311-0. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/264971> (дата обращения: 15.11.2022). — Режим доступа: для авториз. пользователей.

7.2 Дополнительная литература

1. Программные системы статистического анализа. Обнаружение закономерностей в данных с использованием системы R и языка Python : учебное пособие / В. М. Волкова, М. А. Семёнова, Е. С. Четвертакова, С. С. Вожов. — Новосибирск : НГТУ, 2017. — 74 с. — ISBN 978-5-7782-3183-2. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/118287> (дата обращения: 15.11.2022). — Режим доступа: для авториз. Пользователей.
2. Зададаев, С. А. Математика на языке R : учебник / С. А. Зададаев. — Москва : Прометей, 2018. — 324 с. — ISBN 978-5-907003-59-0. — Текст : электронный // Лань : электронно-библиотечная система. — URL:

<https://e.lanbook.com/book/107099> (дата обращения: 15.11.2022). — Режим доступа: для авториз. пользователей.

3. Зыков, С. В. Программирование : учебник и практикум для вузов / С. В. Зыков. — Москва : Издательство Юрайт, 2022. — 320 с. — (Высшее образование). — ISBN 978-5-534-02444-9. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/489754>

7.3 Методические указания, рекомендации и другие материалы к занятиям

1. Харитоновна, А.Е. Хранилища и системы интеллектуального анализа данных: методические указания / А.Е. Харитоновна. – М.: РГАУ-МСХА им. К.А. Тимирязева, 2016. – 25 с.

8. Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины

1. The R Project for Statistical Computing <https://www.r-project.org/>(открытый доступ)
2. The R Project for Statistical Computing <https://www.r-project.org/> (открытый доступ)
3. Анаконда. URL: <https://www.anaconda.com/distribution/>(открытый доступ)
4. Официальный сайт Росстата. URL: <http://www.gks.ru/> (открытый доступ)
5. Официальный сайт Центрального Банка России. URL: <http://www.cbr.ru> (открытый доступ)
6. Bureau of Economic Analysis. URL: <http://www.bea.gov> (открытый доступ)
7. Московская международная валютная биржа. <http://www.micex.ru> (открытый доступ)
8. Официальный сайт Всемирного банка . URL: [http:// www.worldbank.org](http://www.worldbank.org) (открытый доступ)
9. Официальный сайт Министерства финансов РФ. URL: <http://www.minfin.gov.ru> (открытый доступ)
10. Официальный сайт Национального бюро экономических исследований США. URL: [http:// www.nber.org](http://www.nber.org) (открытый доступ)

9. Перечень программного обеспечения и информационных справочных систем

Таблица 9

Перечень программного обеспечения

№ п/п	Наименование раздела учебной дисциплины	Наименование программы	Тип программы	Автор	Год разработки
1	Тема 1. Введение в R Тема 2. Создание набора данных Тема 3. Ввод и импорт данных Тема 4. Графика в R Тема 5. Основы управления данными Тема 6. Числовые и текстовые функции	R	расчётная	r-project	2022

	Тема 7. Создание собственных функций Тема 8. Основные методы статистической обработки данных. Тема 9. Работа с пропущенными данными				
2	Тема 1. Введение в R Тема 2. Создание набора данных Тема 3. Ввод и импорт данных Тема 4. Графика в R Тема 5. Основы управления данными Тема 6. Числовые и текстовые функции Тема 7. Создание собственных функций Тема 8. Основные методы статистической обработки данных. Тема 9. Работа с пропущенными данными	RStudio	расчётная	r-project	2022

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

Таблица 10

Сведения об обеспеченности специализированными аудиториями, кабинетами, лабораториями

Наименование специальных помещений и помещений для самостоятельной работы (№ учебного корпуса, № аудитории)	Оснащённость специальных помещений и помещений для самостоятельной работы
1	2
<i>учебная аудитория для проведения занятий семинарского типа, учебная аудитория для групповых и индивидуальных консультаций, учебная аудитория для текущего контроля и промежуточной аттестации, помещение для самостоятельной работы (2й учебный корпус, 302 ауд.)</i>	<ol style="list-style-type: none"> 1. Системный блок Intel Core Intel Core i3-2100/4096Mb/500Gb/DVD-RW 10 шт. (Инв.№601997, Инв.№601998, Инв.№601999, Инв.№602000, Инв.№602001, Инв.№602002, Инв.№602003, Инв.№602004, Инв.№602005, Инв.№602006) 2. Монитор 10 шт. (без инв. №) - приобретались не за счет средств вуза 3. Шкаф 2 шт. (Инв.№594166, Инв.№594167) 4. Тумба 1 шт. (Инв.№594168) 5. Подвесное крепление к огнетушителю 1 шт. (Инв. № 559528) 6. Огнетушитель порошковый 1 шт. (Инв. №559527) 7. Жалюзи 1 шт. (Инв.№551557) 8. Доска магнитно-маркерная 1 шт. 9. Стол 5 шт. 10. Стол компьютерный 12 шт. 11. Стул офисный 21 шт. 12. Сейф 1 шт. (без Инв.№).
Студенческое общежитие	Комнаты для самоподготовки
ЦНБ имени Н.И. Железнова	Читальный зал

11. Методические рекомендации обучающимся по освоению дисциплины

Приступая к изучению дисциплины «Обработка данных в среде R», студенты должны ознакомиться с учебной программой, учебной, научной и мето-

дической литературой, имеющейся в библиотеке РГАУ-МСХА им. К.А. Тимирязева, получить в библиотеке рекомендованные учебники и учебно-методические пособия, завести новую тетрадь для работы с первоисточниками.

В ходе подготовки к практическим занятиям изучить основную литературу, ознакомиться с дополнительной литературой в соответствии с поставленной задачей. При этом учесть рекомендации преподавателя и требования учебной программы. Необходимо дорабатывать свой конспект, делая в нем соответствующие записи из литературы, рекомендованной преподавателем и предусмотренной учебной программой.

При подготовке к зачету повторять пройденный материал в строгом соответствии с учебной программой. Использовать конспекты и литературу, рекомендованную преподавателем. Обратит особое внимание на темы учебных занятий, пропущенных студентом по разным причинам. При необходимости обратиться за консультацией и методической помощью к преподавателю.

Виды и формы отработки пропущенных занятий

Студент, пропустивший занятия обязан самостоятельно выполнить сообщение (презентацию), рассмотренную на практическом занятии и подготовиться по контрольным вопросам к защите работы в рамках часов консультаций.

12. Методические рекомендации преподавателям по организации обучения по дисциплине

Курс обработки данных в среде R должен давать не абстрактно-формальные, а прикладные знания. Данная цель может быть реализована только при условии соблюдения в учебных планах преемственности учебных дисциплин. Базовые знания для изучения дисциплины дают такие предметы, как экономическая теория, информатика.

Преподаватель должен указывать, в какой последовательности следует изучать материал дисциплины, обращать внимание на особенности изучения отдельных тем и разделов, помогать отбирать наиболее важные и необходимые сведения из учебных пособий, а также давать объяснения вопросам программы курса, которые обычно вызывают затруднения. При этом преподавателю необходимо учитывать следующие моменты:

1. Не следует перегружать студентов творческими заданиями.
2. Чередовать творческую работу на занятиях с заданиями во внеаудиторное время.
3. Давать студентам четкий инструктаж по выполнению самостоятельных заданий: цель задания; условия выполнения; объем; сроки; требования к оформлению.
4. Осуществлять текущий учет и контроль за самостоятельной работой.
5. Давать оценку обобщать уровень усвоения навыков самостоятельной, творческой работы.

Программу разработал (и):

Харитонов А.Е., к.э.н., доцент



(подпись)

РЕЦЕНЗИЯ

на рабочую программу дисциплины ФТД.02 «Обработка данных в среде R»
ОПОП ВО по направлению 09.03.02 «Информационные системы и технологии»,
направленность «Большие данные и машинное обучение (Machine Learning & Big
Data)»

(квалификация выпускника – бакалавр)

Коломеева Елена Сергеевна, доцент кафедры финансов ФГБОУ ВО г. Москвы «РГАУ-МСХА имени К.А. Тимирязева», кандидатом экономических наук (далее по тексту рецензент), проведена рецензия рабочей программы дисциплины «Обработка данных в среде R» ОПОП ВО по направлению 09.03.02 «Информационные системы и технологии», направленность «Большие данные и машинное обучение (Machine Learning & Big Data)» (бакалавриат) разработанной в ФГБОУ ВО «Российский государственный аграрный университет – МСХА имени К.А. Тимирязева», на кафедре статистики и кибернетики (разработчик – Харитоновна Анна Евгеньевна, кандидат экономических наук, доцент кафедры статистики и кибернетики).

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

1. Предъявленная рабочая программа дисциплины «Обработка данных в среде R» (далее по тексту Программа) соответствует требованиям ФГОС ВО по направлению 09.03.02 «Информационные системы и технологии». Программа содержит все основные разделы, соответствует требованиям к нормативно-методическим документам.

1. Представленная в Программе актуальность учебной дисциплины в рамках реализации ОПОП ВО не подлежит сомнению – дисциплина относится к факультативным дисциплинам части, формируемой участниками образовательных отношений учебного цикла учебного цикла – ФТД.

2. Представленные в Программе цели дисциплины соответствуют требованиям ФГОС ВО направления 09.03.02 «Информационные системы и технологии».

3. В соответствии с Программой за дисциплиной «Обработка данных в среде R» закреплено 2 компетенции (3 индикатора). Дисциплина «Обработка данных в среде R» и представленная Программа способна реализовать их в объявленных требованиях.

4. Общая трудоёмкость дисциплины «Обработка данных в среде R» составляет 4 зачётных единицы (144 часов/из них практическая подготовка 8 ч.).

5. Информация о взаимосвязи изучаемых дисциплин и вопросам исключения дублирования в содержании дисциплин соответствует действительности. Дисциплина «Обработка данных в среде R» взаимосвязана с другими дисциплинами ОПОП ВО и Учебного плана по направлению 09.03.02 «Информационные системы и технологии» и возможность дублирования в содержании отсутствует.

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

7. Программа дисциплины «Обработка данных в среде R» предполагает 2 часа занятий в интерактивной форме.

8. Виды, содержание и трудоёмкость самостоятельной работы студентов, представленные в Программе, соответствуют требованиям к подготовке выпускников, содержащимся во ФГОС ВО направления 09.03.02 «Информационные системы и технологии».

9. Представленные и описанные в Программе формы текущей оценки знаний (опрос, как в форме обсуждения отдельных вопросов, так и участие в деловых играх), соответствуют специфике дисциплины и требованиям к выпускникам.

Форма промежуточного контроля знаний студентов, предусмотренная Программой, осуществляется в форме зачета, что соответствует статусу дисциплины, как факультативной дисциплины учебного цикла – ФТД ФГОС ВО направления 09.03.02 «Информационные системы и технологии».

10. Формы оценки знаний, представленные в Программе, соответствуют специфике дисциплины и требованиям к выпускникам.

11. Учебно-методическое обеспечение дисциплины представлено: основной литературой – 4 источника (базовый учебник), дополнительной литературой – 3 наименования, Интернет-ресурсы – 10 источников и соответствует требованиям ФГОС ВО направления 09.03.02 «Информационные системы и технологии».

12. Материально-техническое обеспечение дисциплины соответствует специфике дисциплины «Обработка данных в среде R» и обеспечивает использование современных образовательных, в том числе интерактивных методов обучения.

13. Методические рекомендации студентам и методические рекомендации преподавателям по организации обучения по дисциплине дают представление о специфике обучения по дисциплине «Обработка данных в среде R».

ОБЩИЕ ВЫВОДЫ

На основании проведенного рецензирования можно сделать заключение, что характер, структура и содержание рабочей программы дисциплины «Обработка данных в среде R» ОПОП ВО по направлению 09.03.02 «Информационные системы и технологии», направленность «**Большие данные и машинное обучение (Machine Learning & Big Data)**» (квалификация выпускника – бакалавр), разработанная Харитоновой А. Е., к.э.н., доцентом кафедры статистики и кибернетики, соответствует требованиям ФГОС ВО, современным требованиям экономики, рынка труда и позволит при её реализации успешно обеспечить формирование заявленных компетенций.

Рецензент: Коломеева Е.С., доцент кафедры финансов ФГБОУ ВО «Российский государственный аграрный университет – МСХА имени К.А. Тимирязева», кандидат экономических наук _____ « 26 » _____ 2022 г.

(подпись)