

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

**федеральное государственное автономное
образовательное учреждение высшего образования
«Национальный исследовательский Нижегородский государственный университет
им. Н.И. Лобачевского»**

Факультет социальных наук

УТВЕРЖДЕНО
решением Ученого совета ННГУ
протокол № 6 от 31.05.2023 г.

Рабочая программа дисциплины

Анализ и визуализация данных в R

Уровень высшего образования
Бакалавриат

Направление подготовки / специальность
39.03.01 - Социология

Направленность образовательной программы
Социальная теория и комплексный анализ данных

Форма обучения
очная

г. Нижний Новгород

2023 год начала подготовки

1. Место дисциплины в структуре ОПОП

Дисциплина Б1.В.17 Анализ и визуализация данных в R относится к части, формируемой участниками образовательных отношений образовательной программы.

2. Планируемые результаты обучения по дисциплине, соотнесенные с планируемыми результатами освоения образовательной программы (компетенциями и индикаторами достижения компетенций)

Формируемые компетенции (код, содержание компетенции)	Планируемые результаты обучения по дисциплине (модулю), в соответствии с индикатором достижения компетенции		Наименование оценочного средства	
	Индикатор достижения компетенции (код, содержание индикатора)	Результаты обучения по дисциплине	Для текущего контроля успеваемости	Для промежуточной аттестации
ПК-4: Способен презентовать результаты социологического исследования с учетом выбора целевой аудитории	<p>ПК-4.1: Знать теоретические основы современных методов машинного обучения в сборе и обработке данных.</p> <p>ПК-4.2: Уметь применять методы машинного обучения (обработка текста, аудио и видео материалов).</p> <p>ПК-4.3: Владеть навыком отбора и применения необходимых современных методов сбора и обработки данных для изучения институтов, процессов, управления.</p>	<p>ПК-4.1: Знать:</p> <p>1. Методы сбора данных для их последующего анализа в R.</p> <p>2. Форматы данных для их последующего анализа в R.</p> <p>ПК-4.2: Уметь:</p> <p>1. Обрабатывать данные с помощью математической статистики в R.</p> <p>2. Анализировать данные с помощью математической статистики в R.</p> <p>ПК-4.3: Владеть методами и средствами математической статистики и анализа данных в R.</p>	Кейс-задание	<p>Зачёт:</p> <p>Контрольные вопросы</p> <p>Тест</p> <p>Задания</p>

3. Структура и содержание дисциплины

3.1 Трудоемкость дисциплины

	очная
Общая трудоемкость, з.е.	3
Часов по учебному плану	108
в том числе	
аудиторные занятия (контактная работа):	
- занятия лекционного типа	16

- занятия семинарского типа (практические занятия / лабораторные работы)	16
- КСР	1
самостоятельная работа	75
Промежуточная аттестация	0 зачёт

3.2. Содержание дисциплины

(структурированное по темам (разделам) с указанием отведенного на них количества академических часов и виды учебных занятий)

Наименование и краткое содержание разделов и тем дисциплины	Всего (часы)	в том числе			
		Контактная работа (работа во взаимодействии с преподавателем), часы из них			Самостоятельная работа обучающегося, часы
		Занятия лекционного типа	Занятия семинарского типа (практические занятия/ лабора торные работы), часы	Всего	
	о ф о	о ф о	о ф о	о ф о	о ф о
Тема 1. Основы синтаксиса и работа с фреймами данных	14	2	2	4	10
Тема 2. Описательная статистика	14	2	2	4	10
Тема 3. Визуализация данных	15	2	2	4	11
Тема 4. Анализ номинативных данных	15	2	2	4	11
Тема 5. Сравнение двух групп	15	2	2	4	11
Тема 6. Корреляционный анализ	17	3	3	6	11
Тема 7. Дисперсионный анализ	17	3	3	6	11
Аттестация	0				
КСР	1				1
Итого	108	16	16	33	75

4. Учебно-методическое обеспечение самостоятельной работы обучающихся

По дисциплине предусмотрено выполнение самостоятельной работы в виде анализа конкретного случая и попытки решения проблемы, описанной в кейсе.

По теме кейса студентом самостоятельно собирается материал, обрабатывается, оформляется в виде письменной работы и презентации. Работы обсуждаются на практических занятиях. В ходе группового обсуждения рассматривается та или иная проблема виртуальной реальности. На всех этапах выполнения работы преподаватель оказывает консультативную помощь и методическое обеспечение.

5. Фонд оценочных средств для текущего контроля успеваемости и промежуточной аттестации по дисциплине (модулю)

5.1 Типовые задания, необходимые для оценки результатов обучения при проведении текущего контроля успеваемости с указанием критериев их оценивания:

5.1.1 Типовые задания (оценочное средство - Кейс-задание) для оценки сформированности компетенции ПК-4

1. Как выбрать подходящий графический инструмент для визуализации конкретного типа данных?
2. Как эффективно представить многомерные данные с помощью графиков в RStudio?
3. Как обрабатывать и визуализировать временные ряды данных в RStudio?

Критерии оценивания (оценочное средство - Кейс-задание)

Оценка	Критерии оценивания
зачтено	Минимально допустимый уровень знаний. Допущено много негрубых ошибки. Продемонстрированы основные умения. Решены типовые задачи с негрубыми ошибками. Выполнены все задания, но не в полном объеме. Имеется минимальный набор навыков для решения стандартных задач с некоторыми недочетами
не зачтено	Уровень знаний ниже минимальных требований. Имели место грубые ошибки. При решении стандартных задач не продемонстрированы основные умения. Имели место грубые ошибки. При решении стандартных задач не продемонстрированы базовые навыки. Имели место грубые ошибки.

5.2. Описание шкал оценивания результатов обучения по дисциплине при промежуточной аттестации

Шкала оценивания сформированности компетенций

Уровень сформированности компетенций (индикатора достижения компетенций)	плохо	неудовлетворительно	удовлетворительно	хорошо	очень хорошо	отлично	превосходно
	не зачтено		зачтено				
<u>Знания</u>	Отсутствие знаний теоретического материала. Невозможность оценить полноту знаний вследствие отказа обучающегося от ответа	Уровень знаний ниже минимальных требований. Имели место грубые ошибки	Минимально допустимый уровень знаний. Допущено много негрубых ошибок	Уровень знаний в объеме, соответствующем программе подготовки. Допущено несколько негрубых ошибок	Уровень знаний в объеме, соответствующем программе подготовки. Допущено несколько несущественных ошибок	Уровень знаний в объеме, соответствующем программе подготовки. Ошибок нет.	Уровень знаний в объеме, превышающем программу подготовки.

<u>Умения</u>	Отсутствие минимальных умений. Невозможность оценить наличие умений вследствие отказа обучающегося от ответа	При решении стандартных задач не продемонстрированы основные умения. Имели место грубые ошибки	Продемонстрированы основные умения. Решены типовые задачи с негрубыми ошибками. Выполнены все задания, но не в полном объеме	Продемонстрированы все основные умения. Решены все основные задачи с негрубыми ошибками. Выполнены все задания в полном объеме, но некоторые с недочетами	Продемонстрированы все основные умения. Решены все основные задачи. Выполнены все задания в полном объеме, но некоторые с недочетами	Продемонстрированы все основные умения. Решены все основные задачи с отдельным и несущественными недочетами, выполнены все задания в полном объеме	Продемонстрированы все основные умения. Решены все основные задачи. Выполнены все задания, в полном объеме без недочетов
<u>Навыки</u>	Отсутствие базовых навыков. Невозможность оценить наличие навыков вследствие отказа обучающегося от ответа	При решении стандартных задач не продемонстрированы базовые навыки. Имели место грубые ошибки	Имеется минимальный набор навыков для решения стандартных задач с некоторым и недочетами	Продемонстрированы базовые навыки при решении стандартных задач с некоторым и недочетами	Продемонстрированы базовые навыки при решении стандартных задач без ошибок и недочетов	Продемонстрированы навыки при решении нестандартных задач без ошибок и недочетов	Продемонстрирован творческий подход к решению нестандартных задач

Шкала оценивания при промежуточной аттестации

Оценка		Уровень подготовки
зачтено	превосходно	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «превосходно», продемонстрированы знания, умения, владения по соответствующим компетенциям на уровне выше предусмотренного программой
	отлично	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «отлично».
	очень хорошо	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «очень хорошо»
	хорошо	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «хорошо».
	удовлетворительно	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «удовлетворительно», при этом хотя бы одна компетенция сформирована на уровне «удовлетворительно»
не зачтено	неудовлетворительно	Хотя бы одна компетенция сформирована на уровне «неудовлетворительно».
	плохо	Хотя бы одна компетенция сформирована на уровне «плохо»

5.3 Типовые контрольные задания или иные материалы, необходимые для оценки результатов обучения на промежуточной аттестации

5.3.1 Типовые задания, выносимые на промежуточную аттестацию:

Оценочное средство - Контрольные вопросы

Зачёт

Критерии оценивания (Контрольные вопросы - Зачёт)

Оценка	Критерии оценивания
зачтено	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «удовлетворительно», при этом хотя бы одна компетенция сформирована на уровне «удовлетворительно»
не зачтено	Хотя бы одна компетенция сформирована на уровне «неудовлетворительно», ни одна из компетенций не сформирована на уровне «плохо»

Типовые задания (Контрольные вопросы - Зачёт) для оценки сформированности компетенции ПК-4 (Способен презентовать результаты социологического исследования с учетом выбора целевой аудитории)

1. Что такое RStudio и для чего он используется?
2. Как создать новый скрипт в RStudio?
3. Как импортировать данные из CSV-файла в RStudio?
4. Каковы основные структуры данных в RStudio?
5. Как создать вектор в RStudio?
6. Как создать матрицу в RStudio?
7. Как создать факторный объект в RStudio?
8. Как создать список в RStudio?
9. Как создать датафрейм в RStudio?
10. Как прочитать данные из базы данных в RStudio?
11. Как получить общую информацию о данных в RStudio?
12. Как вывести первые/последние строки данных в RStudio?
13. Как выбрать определенные столбцы из датафрейма в RStudio?
14. Как фильтровать данные в RStudio по определенным условиям?
15. Как сортировать данные в RStudio?
16. Как добавить новый столбец в датафрейм в RStudio?
17. Как изменить значения в определенных ячейках датафрейма в RStudio?
18. Как удалить столбцы из датафрейма в RStudio?
19. Как объединить два датафрейма по общим столбцам в RStudio?
20. Как выполнить простую статистическую сводку данных в RStudio?
21. Как рассчитать сумму/среднее/минимум/максимум по столбцам данных в RStudio?
22. Как выполнить группировку данных по определенному столбцу в RStudio?
23. Как провести базовую визуализацию данных в RStudio?
24. Как построить гистограмму данных в RStudio?
25. Как построить диаграмму рассеяния данных в RStudio?
26. Как построить круговую диаграмму в RStudio?
27. Как создать собственный график в RStudio?
28. Как сохранить график в файл в RStudio?
29. Как выполнить простую линейную регрессию в RStudio?
30. Как оценить статистическую значимость коэффициентов регрессии в RStudio?
31. Как рассчитать корреляцию между двумя переменными в RStudio?

32. Как выполнить т-тест на различие средних значений двух групп данных в RStudio?
33. Как выполнить дисперсионный анализ (ANOVA) в RStudio?
34. Как выполнить кластерный анализ данных в RStudio?
35. Как выполнить факторный анализ данных в RStudio?
36. Как выполнить анализ главных компонент (PCA) в RStudio?
37. Как выполнить временные ряды анализа данных в RStudio?
38. Как выполнить классификацию данных с использованием метода k ближайших соседей в RStudio?
39. Как выполнить классификацию данных с использованием метода логистической регрессии в RStudio?
40. Как выполнить классификацию данных с использованием метода дерева решений в RStudio?
41. Как выполнить кросс-валидацию модели в RStudio?
42. Как выполнить настройку гиперпараметров модели с использованием метода сетки параметров (grid search) в RStudio?
43. Как выполнить анализ текстовых данных в RStudio?
44. Как выполнить анализ социальных сетей (SNA) в RStudio?
45. Как выполнить анализ временных рядов с использованием метода скользящего среднего в RStudio?
46. Как выполнить анализ прогнозирования временных рядов с использованием модели ARIMA в RStudio?
47. Как выполнить анализ прогнозирования временных рядов с использованием модели GARCH в RStudio?
48. Как выполнить анализ прогнозирования временных рядов с использованием модели Prophet в RStudio?
49. Как выполнить анализ выживаемости (survival analysis) в RStudio?
50. Как выполнить анализ машинного обучения с использованием метода случайного леса (random forest) в RStudio?
51. Как выполнить анализ машинного обучения с использованием метода опорных векторов (SVM) в RStudio?
52. Как выполнить анализ машинного обучения с использованием метода градиентного бустинга (gradient boosting) в RStudio?
53. Как выполнить анализ машинного обучения с использованием метода нейронных сетей (neural networks) в RStudio?
54. Как выполнить анализ машинного обучения с использованием метода кластеризации K-средних (K-means) в RStudio?
55. Как выполнить анализ машинного обучения с использованием метода гауссовской смеси (GMM) в RStudio?
56. Как выполнить анализ машинного обучения с использованием метода рекомендательных систем (collaborative filtering) в RStudio?
57. Как выполнить анализ машинного обучения с использованием метода нейронных сетей глубокого обучения (deep learning) в RStudio?
58. Как выполнить анализ машинного обучения с использованием метода ассоциативных правил (association rules) в RStudio?
59. Как выполнить анализ машинного обучения с использованием метода генетических алгоритмов (genetic algorithms) в RStudio?
60. Как выполнить визуализацию пространственных данных с использованием географических информационных систем (GIS) в RStudio?

Оценочное средство - Тест

Зачёт

Критерии оценивания (Тест - Зачёт)

Оценка	Критерии оценивания

Оценка	Критерии оценивания
зачтено	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «удовлетворительно», при этом хотя бы одна компетенция сформирована на уровне «удовлетворительно»
не зачтено	Хотя бы одна компетенция сформирована на уровне «неудовлетворительно», ни одна из компетенций не сформирована на уровне «плохо»

Типовые задания (Тест - Зачёт) для оценки сформированности компетенции ПК-4

(Способен презентовать результаты социологического исследования с учетом выбора целевой аудитории)

Инструкция: Выберите один вариант ответа.

1. Какая функция используется для чтения данных из CSV-файла в RStudio?

1. read.table()
2. read.csv()
3. read.delim()
4. read.xlsx()

2. Какая функция используется для просмотра структуры данных в RStudio?

1. head()
2. str()
3. summary()
4. view()

3. Какая функция используется для фильтрации данных в RStudio?

1. filter()
2. subset()
3. select()
4. arrange()

Оценочное средство - Задания

Зачёт

Критерии оценивания (Задания - Зачёт)

Оценка	Критерии оценивания
зачтено	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «удовлетворительно», при этом хотя бы одна компетенция сформирована на уровне «удовлетворительно»
не зачтено	Хотя бы одна компетенция сформирована на уровне «неудовлетворительно», ни одна из компетенций не сформирована на уровне «плохо»

Типовые задания (Задания - Зачёт) для оценки сформированности компетенции ПК-4

(Способен презентовать результаты социологического исследования с учетом выбора целевой аудитории)

Задание 1

Импортировать данные из CSV файла и вывести первые 5 строк.

Задание 2

Посчитать среднее значение столбца "Возраст" в импортированных данных.

Задание 3

Создать гистограмму для переменной "Оценка" и добавить подписи осей.

6. Учебно-методическое и информационное обеспечение дисциплины (модуля)

Основная литература:

1. Демидова О. А. Эконометрика : учебник и практикум / О. А. Демидова, Д. И. Малахов. - Москва : Юрайт, 2023. - 334 с. - (Высшее образование). - ISBN 978-5-534-00625-4. - Текст : электронный // ЭБС "Юрайт"., <https://e-lib.unn.ru/MegaPro/UserEntry?Action=FindDocs&ids=842844&idb=0>.
2. Крутских В. В. Моделирование в LabVIEW : учебное пособие / В. В. Крутских. - Москва : Юрайт, 2023. - 171 с. - (Высшее образование). - ISBN 978-5-534-13681-4. - Текст : электронный // ЭБС "Юрайт"., <https://e-lib.unn.ru/MegaPro/UserEntry?Action=FindDocs&ids=847701&idb=0>.
3. Дубина И. Н. Основы математического моделирования социально-экономических процессов : учебник и практикум / И. Н. Дубина. - Москва : Юрайт, 2023. - 349 с. - (Высшее образование). - ISBN 978-5-534-00501-1. - Текст : электронный // ЭБС "Юрайт"., <https://e-lib.unn.ru/MegaPro/UserEntry?Action=FindDocs&ids=848852&idb=0>.
4. Информационные системы и технологии в экономике и управлении в 2 ч. Часть 1 : учебник / ответственный редактор В. В. Трофимов. - 5-е изд. ; пер. и доп. - Москва : Юрайт, 2023. - 375 с. - (Высшее образование). - ISBN 978-5-534-09090-1. - Текст : электронный // ЭБС "Юрайт"., <https://e-lib.unn.ru/MegaPro/UserEntry?Action=FindDocs&ids=847797&idb=0>.

Дополнительная литература:

1. Анализ данных : учебник / В. С. Мхитарян [и др.] ; под редакцией В. С. Мхитаряна. - Москва : Юрайт, 2023. - 490 с. - (Высшее образование). - ISBN 978-5-534-00616-2. - Текст : электронный // ЭБС "Юрайт"., <https://e-lib.unn.ru/MegaPro/UserEntry?Action=FindDocs&ids=847079&idb=0>.
2. Миркин Б. Г. Введение в анализ данных : учебник и практикум / Б. Г. Миркин. - Москва : Юрайт, 2023. - 174 с. - (Высшее образование). - ISBN 978-5-9916-5009-0. - Текст : электронный // ЭБС "Юрайт"., <https://e-lib.unn.ru/MegaPro/UserEntry?Action=FindDocs&ids=843081&idb=0>.

Программное обеспечение и Интернет-ресурсы (в соответствии с содержанием дисциплины):

Интернет-ресурсы:

1. <https://stepik.org/course/Анализ-данных-в-R-129/>

7. Материально-техническое обеспечение дисциплины (модуля)

Учебные аудитории для проведения учебных занятий, предусмотренных образовательной программой, оснащены мультимедийным оборудованием (проектор, экран), техническими средствами обучения, компьютерами.

Помещения для самостоятельной работы обучающихся оснащены компьютерной техникой с возможностью подключения к сети "Интернет" и обеспечены доступом в электронную информационно-образовательную среду.

Программа составлена в соответствии с требованиями ОС ННГУ по направлению 39.03.01 - Социология.

Автор(ы): Демарева Валерия Алексеевна, кандидат психологических наук.

Заведующий кафедрой: Голубин Роман Викторович, кандидат исторических наук.

Программа одобрена на заседании методической комиссии от 11.11.2022, протокол № 3.