

Документ подписан простой электронной подписью

Информация о владельце:

ФИО: Макаренко Елена Николаевна

Должность: Ректор

Дата подписания: 10.12.2024 15:06:09

Уникальный программный ключ:

c098bc0c1041cb2a4cf926cf171d6715d99a6ae00adc8e27b55cbe1e2dbd7c78

Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение высшего образования «Ростовский государственный экономический университет (РИНХ)»

УТВЕРЖДАЮ

Начальник

учебно-методического управления

Платонова Т.К.

«25» июня 2024 г.

Рабочая программа дисциплины
Статистический анализ данных в Python

Направление 01.03.05 Статистика
Направленность 01.03.05.01 Анализ больших данных

Для набора 2023 года

Квалификация
Бакалавр

КАФЕДРА Информационных систем и прикладной информатики**Распределение часов дисциплины по семестрам**

Семестр (<Курс>.<Семестр на курсе>)	5 (3.1)		Итого	
	16			
Неделя	16			
Вид занятий	УП	РП	УП	РП
Лекции	10	10	10	10
Лабораторные	12	12	12	12
Итого ауд.	22	22	22	22
Контактная работа	22	22	22	22
Сам. работа	185	185	185	185
Часы на контроль	9	9	9	9
Итого	216	216	216	216

ОСНОВАНИЕ

Учебный план утвержден учёным советом вуза от 25.06.2024 г. протокол № 18.

Программу составил(и): д.э.н., доц., Щербаков С.М.; ст. преп., Шкодина Т.А.

Зав. кафедрой: д.э.н., проф. Щербаков С.М.

Методический совет направления: к.э.н., доцент Андреева О.В.

1. ЦЕЛИ ОСВОЕНИЯ ДИСЦИПЛИНЫ

1.1	развитие знаний и практических навыков анализа статистических данных с использованием алгоритмов, созданных на языке Python.
-----	--

2. ТРЕБОВАНИЯ К РЕЗУЛЬТАТАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ

ПК-1: Способен разрабатывать и реализовывать в виде программного модуля алгоритм решения поставленной теоретической или прикладной задачи

ПК-6: Способен осуществлять поиск статистической информации, ее первичную обработку и подготовку для проведения аналитических исследований, в том числе с использованием технологий больших данных

В результате освоения дисциплины обучающийся должен:

Знать:

методы статистического анализа данных (соотнесено с индикатором ПК-1.1)

методы поиска и первичной обработки статистической информации (соотнесено с индикатором ПК-6.1)

Уметь:

разрабатывать и реализовывать алгоритмы решения задач (соотнесено с индикатором ПК-1.2)

проводить аналитические исследования с использованием технологий больших данных (соотнесено с индикатором ПК-6.2)

Владеть:

навыками реализации алгоритмов задач статистического анализа в Python (соотнесено с индикатором ПК-1.3)

навыками поиска, обработки и анализа статистических данных с использованием технологий больших данных в Python (соотнесено с индикатором ПК-6.3)

3. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

Раздел 1. Методы и алгоритмы аналитики статистических данных

№	Наименование темы / Вид занятия	Семестр / Курс	Часов	Компетенции	Литература
1.1	Тема 1. Методы научных и прикладных исследований в анализе данных. Методы и приемы экономического анализа. Линейные программы. Инструменты анализа данных и прикладное программное обеспечение. Разветвляющиеся вычислительные процессы. / Лек /	5	2	ПК-1, ПК-6	Л2.1, Л2.2, Л2.3
1.2	Тема 2. Математический аппарат (refresher). Введение в модуль NumPy. Основы работы с Pandas. Разведывательный анализ данных. Способы обработки экономической информации в экономическом анализе средствами Python. Организация циклов. Основы машинного обучения в статистическом анализе данных. Организация циклов. / Лек /	5	2	ПК-1, ПК-6	Л2.1, Л2.2, Л2.3
1.3	Тема 3. Работа с текстовыми данными. Сбор данных из открытых источников. Предварительная обработка текстовых данных. Одномерные массивы. Технологии использования статистических данных интернета вещей. Одномерные массивы. Основные принципы организации распределенных систем обработки информации. Двумерные массивы и функции. / Лек /	5	2	ПК-1, ПК-6	Л2.1, Л2.2, Л2.3
1.4	Лабораторное задание 1. Методы научных и прикладных исследований в анализе данных. Лабораторное задание 2. Методы и приемы экономического анализа. Линейные программы. Лабораторное задание 3. Инструменты анализа данных и прикладное программное обеспечение. Разветвляющиеся вычислительные процессы. Лабораторное задание 4. Математический аппарат (refresher). Введение в модуль NumPy. Основы работы с Pandas. Разведывательный анализ данных. Выполнение лабораторных заданий на языке Python с использованием LibreOffice. / Лаб /	5	2	ПК-1, ПК-6	Л2.1, Л2.2, Л2.3

1.5	Лабораторное задание 5. Способы обработки экономической информации в экономическом анализе средствами Python. Организация циклов. Лабораторное задание 6. Основы машинного обучения в статистическом анализе данных. Организация циклов. Лабораторное задание 7. Работа с текстовыми данными. Сбор данных из открытых источников. Предварительная обработка текстовых данных. Одномерные массивы. Лабораторное задание 8. Технологии использования статистических данных интернета вещей. Одномерные массивы. Выполнение лабораторных заданий на языке Python с использованием LibreOffice. / Лаб /	5	2	ПК-1, ПК-6	Л2.1, Л2.2, Л2.3
1.6	Методы научных и прикладных исследований в анализе данных. Линейные программы. / Ср /	5	20	ПК-1, ПК-6	Л2.1, Л2.2, Л2.3
1.7	Инструменты анализа данных и прикладное программное обеспечение. / Ср /	5	20	ПК-1, ПК-6	Л2.1, Л2.2, Л2.3
1.8	Основы работы с Pandas. / Ср /	5	20	ПК-1, ПК-6	Л2.1, Л2.2, Л2.3
1.9	Технологии анализа данных Интернета вещей. / Ср /	5	20	ПК-1, ПК-6	Л2.1, Л2.2, Л2.3
Раздел 2. Инструменты Python для обработки и анализа данных					
№	Наименование темы / Вид занятия	Семестр / Курс	Часов	Компетенции	Литература
2.1	Тема 4. Анализ сетей. Двумерные массивы и функции. Основы программирования на языке Python (refresher). Файлы. Введение в машинное обучение. Модуль sklearn. Задачи классификации и линейные модели. Деревья решений. Случайный лес. Ансамбли моделей. / Лек /	5	2	ПК-1, ПК-6	Л2.1, Л2.2, Л2.3
2.2	Тема 5. Визуализация данных. Представление результатов исследования. Программирование графики. Введение в нейронные сети. Программирование графики. Интерпретация результатов анализа статистических данных. / Лек /	5	2	ПК-1, ПК-6	Л2.1, Л2.2, Л2.3
2.3	Лабораторное задание 9. Основные принципы организации распределенных систем обработки информации. Двумерные массивы и функции. Лабораторное задание 10. Анализ сетей. Двумерные массивы и функции. Лабораторное задание 11. Основы программирования на языке Python (refresher). Файлы. Лабораторное задание 12. Введение в машинное обучение. Модуль sklearn. Задачи классификации и линейные модели. Выполнение лабораторных заданий на языке Python с использованием LibreOffice. / Лаб /	5	4	ПК-1, ПК-6	Л2.1, Л2.2, Л2.3
2.4	Лабораторное задание 13. Деревья решений. Случайный лес. Ансамбли моделей. Лабораторное задание 14. Визуализация данных. Представление результатов исследования. Программирование графики. Лабораторное задание 15. Введение в нейронные сети. Программирование графики. Лабораторное задание 16. Интерпретация результатов анализа статистических данных. Выполнение лабораторных заданий на языке Python с использованием LibreOffice. / Лаб /	5	4	ПК-1, ПК-6	Л2.1, Л2.2, Л2.3
2.5	Распределенные системы обработки данных и анализ сетей. / Ср /	5	20	ПК-1, ПК-6	Л2.1, Л2.2, Л2.3
2.6	Введение в машинное обучение. / Ср /	5	20	ПК-1, ПК-6	Л2.1, Л2.2, Л2.3
2.7	Деревья решений. / Ср /	5	20	ПК-1, ПК-6	Л2.1, Л2.2, Л2.3
2.8	Введение в нейронные сети. / Ср /	5	20	ПК-1, ПК-6	Л2.1, Л2.2, Л2.3
2.9	Интерпретация данных анализа в Python. / Ср /	5	25	ПК-1, ПК-6	Л2.1, Л2.2, Л2.3
2.10	/ Экзамен /	5	9	ПК-1, ПК-6	Л2.1, Л2.2, Л2.3

4. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

Структура и содержание фонда оценочных средств для проведения текущей и промежуточной аттестации представлены в Приложении 1 к рабочей программе дисциплины.

5. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

5.2. Дополнительная литература

	Авторы,	Заглавие	Издательство, год	Колич-во
Л2.1	Емельянов А. А.	Прикладная информатика: журнал	Москва: Синергия ПРЕСС, 2006	https://biblioclub.ru/index.php?page=book&id=120298 неограниченный доступ для зарегистрированных пользователей
Л2.2	Северенс Ч.	Введение в программирование на Python: учебное пособие	Москва: Национальный Открытый Университет «ИНТУИТ», 2016	https://biblioclub.ru/index.php?page=book&id=429184 неограниченный доступ для зарегистрированных пользователей
Л2.3	Минашкин, В. Г., Садовникова, Н. А., Шмойлова, Р. А.	Бизнес-статистика и прогнозирование: учебное пособие	Москва: Евразийский открытый институт, 2010	https://www.iprbookshop.ru/10624.html неограниченный доступ для зарегистрированных пользователей

5.3 Профессиональные базы данных и информационные справочные системы

Национальная электронная библиотека (НЭБ) - <https://rusneb.ru/>

ИСС «КонсультантПлюс»

ИСС «Гарант» <http://www.internet.garant.ru/>

5.4. Перечень программного обеспечения

Операционная система РЕД ОС

LibreOffice

5.5. Учебно-методические материалы для студентов с ограниченными возможностями здоровья

При необходимости по заявлению обучающегося с ограниченными возможностями здоровья учебно-методические материалы предоставляются в формах, адаптированных к ограничениям здоровья и восприятия информации. Для лиц с нарушениями зрения: в форме аудиофайла; в печатной форме увеличенным шрифтом. Для лиц с нарушениями слуха: в форме электронного документа; в печатной форме. Для лиц с нарушениями опорно-двигательного аппарата: в форме электронного документа; в печатной форме.

6. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Помещения для всех видов работ, предусмотренных учебным планом, укомплектованы необходимой специализированной учебной мебелью и техническими средствами обучения:

- столы, стулья;
- персональный компьютер / ноутбук (переносной);
- проектор;
- экран / интерактивная доска.

Помещения для проведения всех видов работ, предусмотренных учебным планом, укомплектованы необходимой специализированной учебной мебелью и техническими средствами обучения. Лабораторные занятия проводятся в компьютерных классах, рабочие места в которых оборудованы необходимыми лицензионными программными средствами и выходом в Интернет.

7. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ

Методические указания по освоению дисциплины представлены в Приложении 2 к рабочей программе дисциплины.

ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

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

1.1. Показатели и критерии оценивания компетенций:

ЗУН, составляющие компетенцию	Показатели оценивания	Критерии оценивания	Средства оценивания
ПК-1: Способен разрабатывать и реализовывать в виде программного модуля алгоритм решения поставленной теоретической или прикладной задачи			
З методы статистического анализа данных	знает основные понятия и определения	полнота и содержательность ответа умение приводить примеры	О – задания для опроса (1-6), Э – вопросы к экзамену (1-17)
У разрабатывать и реализовывать алгоритмы решения задач	выполняет задания, отвечает на вопросы, умеет применять полученные знания	полнота и содержательность ответа умение приводить примеры умение самостоятельно находить решение поставленных задач	ЛЗ – лабораторные задания (1-8)
В навыками реализации алгоритмов задач статистического анализа в Python	проводит обобщенный анализ информации и обработку данных	полнота и содержательность ответа умение приводить примеры умение самостоятельно находить решение поставленных задач	ЛЗ – лабораторные задания (1-8)
ПК-6: Способен осуществлять поиск статистической информации, ее первичную обработку и подготовку для проведения аналитических исследований, в том числе с использованием технологий больших данных			
З методы поиска и первичной обработки статистической информации	знает основные понятия и определения	полнота и содержательность ответа умение приводить примеры	О – задания для опроса (7-12), Э – вопросы к экзамену (18-35)
У проводить аналитические исследования с использованием технологий больших данных	выполняет задания, отвечает на вопросы, умеет применять полученные знания	полнота и содержательность ответа умение приводить примеры умение самостоятельно находить решение поставленных задач	ЛЗ – лабораторные задания (9-16)
В навыками поиска, обработки и анализа статистических данных с использованием технологий больших данных в Python	проводит обобщенный анализ информации и обработку данных	полнота и содержательность ответа умение приводить примеры умение самостоятельно находить решение поставленных задач	ЛЗ – лабораторные задания (9-16)

1.2 Шкалы оценивания:

Текущий контроль успеваемости и промежуточная аттестация осуществляется в рамках накопительной балльно-рейтинговой системы в 100-балльной шкале.

- 84-100 баллов (оценка «отлично»);
- 67-83 баллов (оценка «хорошо»);
- 50-66 баллов (оценка «удовлетворительно»);
- 0-49 баллов (оценка «неудовлетворительно»).

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

Вопросы к экзамену

1. Методы научных и прикладных исследований в анализе данных.
2. Методы и приемы экономического анализа.
3. Линейные программы.
4. Инструменты анализа данных и прикладное программное обеспечение.
5. Разветвляющиеся вычислительные процессы.
6. Математический аппарат (refresher).
7. Введение в модуль NumPy.
8. Основы работы с Pandas.
9. Разведывательный анализ данных.
10. Способы обработки экономической информации в экономическом анализе средствами Python.
11. Организация циклов.
12. Основы машинного обучения в статистическом анализе данных.
13. Работа с текстовыми данными.
14. Сбор данных из открытых источников.
15. Предварительная обработка текстовых данных.
16. Одномерные массивы.
17. Технологии использования статистических данных интернета вещей.
18. Основные принципы организации распределенных систем обработки информации.
19. Двумерные массивы и функции.
20. Анализ сетей.
21. Двумерные массивы и функции.
22. Основы программирования на языке Python (refresher).
23. Файлы.
24. Введение в машинное обучение.
25. Модуль sklearn.
26. Задачи классификации и линейные модели.
27. Деревья решений.
28. Случайный лес.
29. Ансамбли моделей.
30. Визуализация данных.
31. Представление результатов исследования.
32. Программирование графики.
33. Введение в нейронные сети.
34. Программирование графики.
35. Интерпретация результатов анализа статистических данных.

Экзаменационное задание включает три вопроса – два теоретических вопроса и одно практико-ориентированное задание из числа приведенных ниже лабораторных заданий.

Критерии оценивания:

- 84-100 баллов (оценка «отлично») – изложенный материал фактически верен, наличие глубоких исчерпывающих знаний в объеме пройденной программы дисциплины в соответствии с поставленными программой курса целями и задачами обучения; правильные, уверенные действия по применению полученных знаний на практике, грамотное и логически стройное изложение материала при ответе, усвоение основной и знакомство с дополнительной литературой;

- 67-83 баллов (оценка «хорошо») – наличие твердых и достаточно полных знаний в объеме пройденной программы дисциплины в соответствии с целями обучения, правильные действия по применению знаний на практике, четкое изложение материала, допускаются отдельные логические и стилистические погрешности, обучающийся усвоил основную литературу, рекомендованную в рабочей программе дисциплины;

- 50-66 баллов (оценка «удовлетворительно») – наличие твердых знаний в объеме пройденного курса в соответствии с целями обучения, изложение ответов с отдельными ошибками, уверенно исправленными после дополнительных вопросов; правильные в целом действия по применению знаний на практике;

- 0-49 баллов (оценка «неудовлетворительно») – ответы не связаны с вопросами, наличие грубых ошибок в ответе, непонимание сущности излагаемого вопроса, неумение применять знания на практике, неуверенность и неточность ответов на дополнительные и наводящие вопросы.

Задания для опроса

Вариант 1

Методы научных и прикладных исследований в анализе данных.

Методы и приемы экономического анализа.

Линейные программы.

Вариант 2

Инструменты анализа данных и прикладное программное обеспечение.

Разветвляющиеся вычислительные процессы.

Математический аппарат (refresher).

Вариант 3

Введение в модуль NumPy.

Основы работы с Pandas.

Разведывательный анализ данных.

Вариант 4

Способы обработки экономической информации в экономическом анализе средствами Python.

Организация циклов.

Основы машинного обучения в статистическом анализе данных.

Вариант 5

Работа с текстовыми данными.

Сбор данных из открытых источников.

Предварительная обработка текстовых данных.

Вариант 6

Одномерные массивы.

Технологии использования статистических данных интернета вещей.

Основные принципы организации распределенных систем обработки информации.

Вариант 7

Двумерные массивы и функции.

Анализ сетей.

Двумерные массивы и функции.

Вариант 8

Основы программирования на языке Python (refresher).

Файлы.

Введение в машинное обучение.

Вариант 9

Модуль sklearn.

Задачи классификации и линейные модели.

Деревья решений.

Вариант 10

Случайный лес.

Ансамбли моделей.

Визуализация данных.

Вариант 11

Представление результатов исследования.

Программирование графики.

Введение в нейронные сети.

Вариант 12

Введение в нейронные сети.

Программирование графики.

Интерпретация результатов анализа статистических данных.

Критерии оценивания (для каждого варианта):

16-20 б. – ответы на все вопросы даны верно;

13-15 б. – один из ответов с неточностями;

10-12 б. – 2 ответа с неточностями;

7-9 б. – 3 ответа с неточностями;

4-6 б. – нет ответа на один вопрос;

0-3 б. – нет ответа на 2 вопроса.

Максимальное количество баллов за опрос – 20.

Лабораторные задания

Лабораторное задание 1.

Методы научных и прикладных исследований в анализе данных.

Выполнение лабораторных заданий на языке Python с использованием LibreOffice.

Лабораторное задание 2.

Методы и приемы экономического анализа. Линейные программы.

Выполнение лабораторных заданий на языке Python с использованием LibreOffice.

Лабораторное задание 3.

Инструменты анализа данных и прикладное программное обеспечение. Разветвляющиеся вычислительные процессы.

Выполнение лабораторных заданий на языке Python с использованием LibreOffice.

Лабораторное задание 4.

Математический аппарат (refresher). Введение в модуль NumPy. Основы работы с Pandas. Разведывательный анализ данных.

Выполнение лабораторных заданий на языке Python с использованием LibreOffice.

Лабораторное задание 5.

Способы обработки экономической информации в экономическом анализе средствами Python. Организация циклов.

Выполнение лабораторных заданий на языке Python с использованием LibreOffice.

Лабораторное задание 6.

Основы машинного обучения в статистическом анализе данных. Организация циклов.

Выполнение лабораторных заданий на языке Python с использованием LibreOffice.

Лабораторное задание 7.

Работа с текстовыми данными. Сбор данных из открытых источников. Предварительная обработка текстовых данных. Одномерные массивы.

Выполнение лабораторных заданий на языке Python с использованием LibreOffice.

Лабораторное задание 8.

Технологии использования статистических данных интернета вещей. Одномерные массивы.

Выполнение лабораторных заданий на языке Python с использованием LibreOffice.

Лабораторное задание 9.

Основные принципы организации распределенных систем обработки информации. Двумерные массивы и функции.

Выполнение лабораторных заданий на языке Python с использованием LibreOffice.

Лабораторное задание 10.

Анализ сетей. Двумерные массивы и функции.

Выполнение лабораторных заданий на языке Python с использованием LibreOffice.

Лабораторное задание 11.

Основы программирования на языке Python (refresher). Файлы.

Выполнение лабораторных заданий на языке Python с использованием LibreOffice.

Лабораторное задание 12.

Введение в машинное обучение. Модуль sklearn. Задачи классификации и линейные модели.

Выполнение лабораторных заданий на языке Python с использованием LibreOffice.

Лабораторное задание 13.

Деревья решений. Случайный лес. Ансамбли моделей.

Выполнение лабораторных заданий на языке Python с использованием LibreOffice.

Лабораторное задание 14.

Визуализация данных. Представление результатов исследования. Программирование графики.

Выполнение лабораторных заданий на языке Python с использованием LibreOffice.

Лабораторное задание 15.

Введение в нейронные сети. Программирование графики.

Выполнение лабораторных заданий на языке Python с использованием LibreOffice.

Лабораторное задание 16.

Интерпретация результатов анализа статистических данных.

Выполнение лабораторных заданий на языке Python с использованием LibreOffice.

Критерии оценивания (для каждого задания):

4-5 б. – задание выполнено верно;

2-3 б. – при выполнении задания были допущены несущественные ошибки;

0-1 б. – при выполнении задания были допущены существенные ошибки.

Максимальное количество баллов за лабораторные задания – 80 (16 заданий по 5 баллов).

3. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций

Процедуры оценивания включают в себя текущий контроль и промежуточную аттестацию.

Текущий контроль успеваемости проводится с использованием оценочных средств, представленных в п. 2 данного приложения. Результаты текущего контроля доводятся до сведения студентов до промежуточной аттестации.

Промежуточная аттестация проводится в форме экзамена.

Экзамен проводится по расписанию промежуточной аттестации. Количество вопросов в задании – 3 (два теоретических вопроса и одно практико-ориентированное задание). Объявление результатов производится в день экзамена. Результаты аттестации заносятся в ведомость и зачетную книжку студента. Студенты, не прошедшие промежуточную аттестацию по графику, должны ликвидировать задолженность в установленном порядке.

МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ

Учебным планом предусмотрены следующие виды занятий:

- лабораторные занятия.

В ходе лабораторных занятий развиваются навыки практической работы.

При подготовке к лабораторным занятиям каждый студент должен:

– изучить рекомендованную учебную литературу;

– подготовить ответы на все вопросы по изучаемой теме.

В процессе подготовки к лабораторным занятиям студенты могут воспользоваться консультациями преподавателя.

Вопросы, не рассмотренные на лабораторных занятиях, должны быть изучены студентами в ходе самостоятельной работы. Контроль самостоятельной работы студентов над учебной программой курса осуществляется в ходе занятий методом опроса и выполнения лабораторных заданий. В ходе самостоятельной работы каждый студент обязан прочитать основную и по возможности дополнительную литературу по изучаемой теме, сделать выписки из рекомендованных первоисточников, выделить непонятные термины, найти их значение в энциклопедических словарях.

Студент должен готовиться к предстоящему лабораторному занятию по всем обозначенным в рабочей программе дисциплины вопросам.

Для подготовки к занятиям, текущему контролю и промежуточной аттестации студенты могут воспользоваться электронно-библиотечными системами. Также обучающиеся могут взять на дом необходимую литературу на абонементе университетской библиотеки или воспользоваться читальными залами.