

Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Хоружий Людмила Ивановна
Должность: Директор института экономики и управления АПК
Дата подписания: 07.02.2024 16:28:33
Уникальный программный ключ:
1e90b132d9b04dce67585160b015dddf2cb1e6a9



УТВЕРЖДАЮ:

Директор Института экономики и
управления АПК

Л.И. Хоружий
2023 г.

Лист актуализации рабочей программы дисциплины «Б1.В.24 Технологии работы с открытыми данными»

Направление: 09.03.03 «Прикладная информатика»

Направленность: Прикладная информатика в экономике

Курс 3

Семестр 5


1. В рабочую программу не вносятся изменения. Программа актуализирована для 2023 г. начала подготовки.

2. Программа будет распространена при организации учебного процесса на направленность (профиль): ИТ-решения для бизнеса.

Разработчик (и): Демичев В.В., канд. экон. наук, доцент


Невзоров А.С., ассистент

(ФИО, ученая степень, ученое звание)



(подпись)

«28» августа 2023 г.

Рабочая программа пересмотрена и одобрена на заседании кафедры статистики и кибернетики протокол № 11 от «28» августа 2023 г.

И.о. заведующего кафедрой статистики и кибернетики  А.В. Уколова

Лист актуализации принят на хранение:

И.о. заведующего кафедрой статистики и кибернетики  А.В. Уколова

Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Хоружий Людмила Ивановна
Должность: Директор института экономики и управления АПК
Дата подписания: 15.07.2023 19:28:57
Уникальный программный ключ:
1e90b132d9b04dceb7585160b015dadd12cb1e6a9



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 управления АПК

Л.И. Хоружий
“ 15 ” июля 2022 г.



РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ
Б1.В.25 Технологии работы с открытыми данными

для подготовки бакалавров

ФГОС ВО

Направление: 09.03.03 «Прикладная информатика»

Направленность: Прикладная информатика в экономике

Курс 2

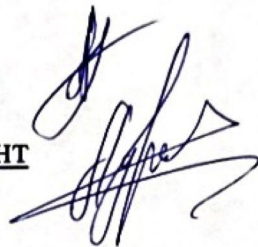
Семестр 4

Форма обучения очная

Год начала подготовки 2022

Москва, 2022

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



«26» 08 2022 г.

Ульянкин А.Е., ассистент
(ФИО, ученая степень, ученое звание)

«26» 08 2022 г.

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



(подпись)
«26» 08 2022 г.

Программа составлена в соответствии с требованиями ФГОС ВО по направлению подготовки 09.03.03 «Прикладная информатика» и учебного плана

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

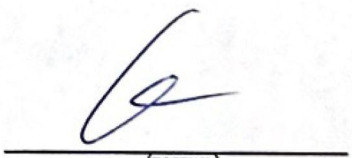
И.о. зав. кафедрой Уколова А.В., к.э.н., доцент
(ФИО, ученая степень, ученое звание)



(подпись)
«26» 08 2022 г.

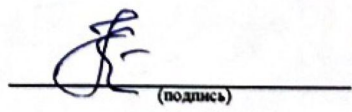
Согласовано:

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



(подпись)
«26» 08 2022 г.

И.о. зав. выпускающей кафедрой
прикладной информатики
Худякова Е.В., д.э.н., проф.
(ФИО, ученая степень, ученое звание)



(подпись)
«26» 08 2022 г.

Заведующий отделом комплектования ЦНБ



(подпись)

СОДЕРЖАНИЕ

АННОТАЦИЯ	4
1. ЦЕЛЬ ОСВОЕНИЯ ДИСЦИПЛИНЫ	5
2. МЕСТО ДИСЦИПЛИНЫ В УЧЕБНОМ ПРОЦЕССЕ	5
3. ПЕРЕЧЕНЬ ПЛАНИРУЕМЫХ РЕЗУЛЬТАТОВ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ, СООТНЕСЕННЫХ С ПЛАНИРУЕМЫМИ РЕЗУЛЬТАТАМИ ОСВОЕНИЯ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ	5
4. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ.....	5
4.1 РАСПРЕДЕЛЕНИЕ ТРУДОЁМКОСТИ ДИСЦИПЛИНЫ ПО ВИДАМ РАБОТ	5
ПО СЕМЕСТРАМ	5
4.2 СОДЕРЖАНИЕ ДИСЦИПЛИНЫ	7
4.3 ПРАКТИЧЕСКИЕ ЗАНЯТИЯ	8
5. ОБРАЗОВАТЕЛЬНЫЕ ТЕХНОЛОГИИ.....	10
6. ТЕКУЩИЙ КОНТРОЛЬ УСПЕВАЕМОСТИ И ПРОМЕЖУТОЧНАЯ АТТЕСТАЦИЯ ПО ИТОГАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ.....	10
6.1. ТИПОВЫЕ КОНТРОЛЬНЫЕ ЗАДАНИЯ ИЛИ ИНЫЕ МАТЕРИАЛЫ, НЕОБХОДИМЫЕ ДЛЯ ОЦЕНКИ ЗНАНИЙ, УМЕНИЙ И НАВЫКОВ И (ИЛИ) ОПЫТА ДЕЯТЕЛЬНОСТИ.....	10
6.2. ОПИСАНИЕ ПОКАЗАТЕЛЕЙ И КРИТЕРИЕВ КОНТРОЛЯ УСПЕВАЕМОСТИ, ОПИСАНИЕ ШКАЛ ОЦЕНИВАНИЯ	16
7. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ	16
7.1 Основная литература	17
7.2 Дополнительная литература	17
8. ПЕРЕЧЕНЬ РЕСУРСОВ ИНФОРМАЦИОННО-ТЕЛЕКОММУНИКАЦИОННОЙ СЕТИ «ИНТЕРНЕТ», НЕОБХОДИМЫХ ДЛЯ ОСВОЕНИЯ ДИСЦИПЛИНЫ.....	18
9. ПЕРЕЧЕНЬ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ	18
10. ОПИСАНИЕ МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЙ БАЗЫ, НЕОБХОДИМОЙ ДЛЯ ОСУЩЕСТВЛЕНИЯ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА ПО ДИСЦИПЛИНЕ	19
11. МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ СТУДЕНТАМ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ.....	19
12. МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ ПРЕПОДАВАТЕЛЯМ ПО ОРГАНИЗАЦИИ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ	20

Аннотация

рабочей программы по дисциплине Б1.В.25 «Технологии работы с открытыми данными» для подготовки бакалавров по направлению 09.03.03 «Прикладная информатика» по направленности «Прикладная информатика в экономике»

Цель освоения дисциплины: Целью дисциплины «Технологии работы с открытыми данными» является освоение студентами теоретических и практических знаний и приобретение умений и навыков работы с данными, их поиска, анализа, визуализации посредством использования языков программирования Python и R.

Место дисциплины в учебном плане: дисциплина включена в часть, формируемую участниками образовательных отношений учебного плана по направлению подготовки 09.03.03 «Прикладная информатика».

Требования к результатам освоения дисциплины: в результате освоения дисциплины формируются следующие компетенции (индикаторы): УК-1 (УК-1.1; УК-1.2; УК-1.3); УК-4 (УК-4.1).

Краткое содержание дисциплины:

Введение в данные: источники, типы данных, качество данных. Правовые аспекты открытых данных. Открытые данные в России и за рубежом.

Методология работы с данными. Основные описательные статистики: средняя, мода, медиана. Проверка и подготовка данных. Расчет описательных статистик на Python и R.

Типы данных. Форматы данных: Json, csv, xml. Различия и методы работы и сферы применения. Загрузка данных различных форматов в среду программирования. Библиотека Pandas. Работа с Data Frame.

Визуализация данных. Виды графиков, построение диаграмм в Python и R.

Методы анализа данных в Python и R. Классификация, регрессия, машинное обучение (Дерево решение, случайный лес, логистическая регрессия и др.). Подбор параметров моделей машинного обучения.

Общая трудоемкость дисциплины составляет: 3 зачетные единицы (108 часов).

Промежуточный контроль: зачет.

1. Цель освоения дисциплины

Целью дисциплины «Технологии работы с открытыми данными» является освоение студентами теоретических и практических знаний и приобретение умений и навыков работы с данными, их поиска, анализа, визуализации посредством использования языков программирования Python и R.

2. Место дисциплины в учебном процессе

Дисциплина «Технологии работы с открытыми данными» включена в часть, формируемую участниками образовательных отношений учебного плана. Дисциплина «Технологии работы с открытыми данными» реализуется в соответствии с требованиями ФГОС ВО, ОПОП ВО и Учебного плана по направлению 09.03.03 «Прикладная информатика».

Дисциплина «Технологии работы с открытыми данными» является основополагающей для изучения следующих дисциплин: «Интеллектуальный анализ данных», «Информационные системы и технологии», «Исследование операций и методы оптимизации», «ИТ-системы в экономике» и «Системы поддержки принятия решений».

Рабочая программа дисциплины «Технологии работы с открытыми данными» 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

Изучение данной учебной дисциплины направлено на формирование у обучающихся компетенций, представленных в таблице 1.

4. Структура и содержание дисциплины

4.1 Распределение трудоёмкости дисциплины по видам работ по семестрам

Общая трудоёмкость дисциплины составляет 3 зач.ед. (108 часов), 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-1.1 Знает принципы сбора, отбора и обобщения информации, методики системного подхода для решения профессиональных задач	методики поиска, сбора и обработки информации в сфере профессиональной деятельности	-	-
			УК-1.2 Умеет анализировать и систематизировать разнородные данные, оценивать эффективность процедур анализа проблем и принятия решений в профессиональной	-	применять методики поиска, сбора и обработки информации в сфере профессиональной деятельности, проводить критический анализ информации	-
			УК-1.3 Владеет навыками научного поиска и практической работы с информационными источниками; методами принятия решений.	-	-	методами поиска, сбора и обработки, критического анализа информации в сфере профессиональной деятельности
2.	УК-4	Способен осуществлять деловую коммуникацию в устной и письменной формах на государственном языке Российской Федерации и иностранном(ых) языке(ах)	УК-4.1 Знает принципы построения устного и письменного высказывания на государственном и иностранном языках; требования к деловой устной и письменной	современные средства информационно-коммуникационных технологий, языковой материал для общения в сфере профессиональной деятельности	-	-

Таблица 2

Распределение трудоёмкости дисциплины по видам работ по семестрам

Вид учебной работы	Трудоёмкость	
	час. всего/*	В т.ч. по семестрам № 4
Общая трудоёмкость дисциплины по учебному плану	108	108
1. Контактная работа:	50,25	50,25
Аудиторная работа	50,25	50,25
<i>лекции (Лек)</i>	16	16
<i>практические занятия (ПЗ)</i>	34/0	34/0
<i>контактная работа на промежуточном контроле (КРА)</i>	0,25	0,25
2. Самостоятельная работа (СРС)	57,75	57,75
<i>самостоятельное изучение разделов, самоподготовка (проработка и повторение лекционного материала и материала учебников и учебных пособий, подготовка к практическим занятиям, и т.д.)</i>	48,75	48,75
<i>Подготовка к зачету</i>	9	9
Вид промежуточного контроля:	Зачет	

4.2 Содержание дисциплины

Таблица 3

Тематический план учебной дисциплины

Наименование разделов и тем дисциплин (укрупнённо)	Всего	Аудиторная работа			Внеаудиторная работа СР
		ПЗ/*	Лек	ПКР/*	
Раздел 1. Открытые данные. Правовые аспекты. Источники данных	8	2	2		4
Раздел 2. Методы работы с данными. Описательные статистики	21,75	8	4		9,75
Раздел 3. Типы данных. Загрузка данных. Работа с таблицами	32	10	4		18
Раздел 4. Визуализация данных	14	4	2		8
Раздел 5. Обработка данных	32	10	4		18
Контактная работа на промежуточном контроле (КРА)	0,25			0,25	
Итого по дисциплине	108	34	16	0,25	57,75

Раздел 1. Открытые данные. Правовые аспекты. Источники данных

Основные понятия и термины открытых данных. Особенности работы с данными. Открытые данные и исследования. Источники данных, их виды. Достоверность источников данных. Качество данных. Правовое регулирование открытых данных. Открытые данные в России. Открытые данные за рубежом.

Раздел 2. Методы работы с данными. Описательные статистики

Основные статистики работы с открытыми данными. Ряды распределения. Среднее значение, мода, медиана. Показатели вариации. Расчет описательных статистик на Python. Основы работы с R. Реализация описательных статистик в R.

Раздел 3. Типы данных. Загрузка данных. Работа с таблицами

Основные типы данных, форматы данных. Json, csv, xml и другие. Как с ними работать. Чем отличаются. В каких случаях используются. Загрузка данных в Python и R. Работа с Data Frame. Работа с типом Series.

Раздел 4. Визуализация данных

Системы визуализации данных. Визуализация данных в Python и R. Библиотеки для визуализации данных. Анализ на основе визуального представления.

Раздел 5. Обработка данных

Введение в машинное обучение. Задачи классификации и линейные модели. Обучение с учителем и без учителя. Проблема переобучения. Дерево решений. Случайный лес. Эффективные библиотеки работы с данными на Python и R. Машинное обучение на Python и R.

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
1	Раздел 1. Открытые данные. Правовые аспекты. Источники данных	Практическая работа № 1. Основные термины и понятия открытых данных	УК-1.1 УК-1.2 УК-1.3 УК-4.1	Вопросы дискуссии	2
2	Раздел 2. Методы работы с данными. Описательные статистики	Практическая работа № 2. Построение рядов распределения в Excel, Python, R	УК-1.1 УК-1.2 УК-1.3 УК-4.1	Защита работы	4
		Практическая работа № 3. Описательные статистики: средняя, мода, медиана в Excel, Python, R	УК-1.1 УК-1.2 УК-1.3 УК-4.1	Защита работы	2
		Практическая работа № 4. Показатели вариации в Excel, Python, R	УК-1.1 УК-1.2 УК-1.3 УК-4.1	Защита работы	2
3	Раздел 3. Типы данных. Загрузка данных. Работа с таблицами	Практическая работа № 5. Загрузка данных в Python, работа с типом Series	УК-1.1 УК-1.2 УК-1.3 УК-4.1	Защита работы	2
		Практическая работа № 6. За-	УК-1.1	Защита ра-	4

№ п/п	№ раздела	№ и название практических занятий	Формируемые компетенции (индикаторы)	Вид контрольного мероприятия	Кол-во часов
		грузка данных в Python, работа с Data Frame	УК-1.2 УК-1.3 УК-4.1	боты	
		Практическая работа 7. Загрузка данных в R.	УК-1.1 УК-1.2 УК-1.3 УК-4.1	Защита работы	2
		Практическая работа 8. Обработка данных в Data Frame	УК-1.1 УК-1.2 УК-1.3 УК-4.1	Защита работы	2
4	Раздел 4. Визуализация данных	Практическая работа № 9. Библиотеки для визуализации данных в Python	УК-1.1 УК-1.2 УК-1.3 УК-4.1	Защита работы	2
		Практическая работа № 10. Библиотеки для визуализации данных в R	УК-1.1 УК-1.2 УК-1.3 УК-4.1	Защита работы	2
5	Раздел 5. Обработка данных	Практическая работа № 11. Построение моделей корреляционно-регрессионного анализа	УК-1.1 УК-1.2 УК-1.3 УК-4.1	Защита работы	2
		Практическая работа № 12. Классификация методом машинного обучения «Дерево решений»	УК-1.1 УК-1.2 УК-1.3 УК-4.1	Защита работы	2
		Практическая работа № 13. Другие методы машинного обучения	УК-1.1 УК-1.2 УК-1.3 УК-4.1	Защита работы	2
		Практическая работа № 14. Оптимизация моделей, подбор параметров	УК-1.1 УК-1.2 УК-1.3 УК-4.1	Защита работы	2
		Контрольная работа	УК-1.1 УК-1.2 УК-1.3 УК-4.1	Контрольная работа	2

Таблица 5

Перечень вопросов для самостоятельного изучения дисциплины

№ п/п	№ раздела и темы	Перечень рассматриваемых вопросов для самостоятельного изучения
1.	Раздел 1. Открытые данные. Правовые аспекты. Источники данных	Открытые данные за рубежом. Правовое регулирование открытых данных (отечественный и зарубежный опыт) (УК-1.1, УК-1.2, УК-1.3, УК-4.1)
2.	Раздел 2. Методы работы с данными. Описательные статистики	Обработка данных в Excel: построение рядов распределения, расчет основных статистик, расчет показателей вариации. (УК-1.1, УК-1.2, УК-1.3, УК-4.1)

№ п/п	№ раздела и темы	Перечень рассматриваемых вопросов для самостоятельного изучения
3	Раздел 3. Типы данных. Загрузка данных. Работа с таблицами	Библиотеки для загрузки данных, работа с Data Frame (УК-1.1, УК-1.2, УК-1.3, УК-4.1)
4	Раздел 4. Визуализация данных	Библиотеки для построения графиков и диаграмм (УК-1.1, УК-1.2, УК-1.3, УК-4.1)
5	Раздел 5. Обработка данных	Методы машинного обучения: логистическая регрессия, наивный Байесовский классификатор, и др. (УК-1.1, УК-1.2, УК-1.3, УК-4.1)

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 № 12. Классификация методом машинного обучения «Дерево решений» без учителя	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

Пример контрольной работы

Вариант 1

1. Загрузить данные из файла.
2. Отсортировать вектор по возрастанию. Построить огиву ранжированного ряда. Удалить элемент, стоящий на 4 месте в отсортированной списке. На вторую позицию вставить значение «б». Создать вектор x , состоящий из 20 случайных чисел. Объединить векторы x и y . Проверить соответствует ли он нормальному закону распределения получившийся вектор по критерию Шапиро. Сделать вывод.
3. Создать функцию, которая автоматически строила дискретный ряд распределения.
4. Создать функцию, которая автоматически рассчитывала бы среднюю, моду, медиану в дискретном ряду распределения.
5. Создать функцию, которая автоматически рассчитывала бы показатели вариации в дискретном ряду распределения.
6. Построить модель парной линейной регрессии между показателями Y и X , дать интерпретацию коэффициенту полной регрессии, сделать прогноз на 5 лет вперед.

7. Построить модель машинного обучения по методу «Дерево решений», оценить точность классификации разными метриками.

Вариант2

1. Загрузить данные из файла.
2. Отсортировать вектор по возрастанию. Построить огиву ранжированного ряда. Удалить элемент, стоящий на 8 месте в отсортированном списке. На вторую позицию вставить значение «321». Создать вектор x , состоящий из 15 случайных чисел. Объединить векторы x и y . Проверить соответствует ли он нормальному закону распределения получившийся вектор по критерию Шапиро. Сделать вывод.
3. Создать функцию, которая автоматически строила интервальный ряд распределения.
4. Создать функцию, которая автоматически рассчитывала бы среднюю, моду, медиану в интервальном ряду распределения.
5. Создать функцию, которая автоматически рассчитывала бы показатели вариации в интервальном ряду распределения.
6. Построить модель парной линейной регрессии между показателями Y и X , дать интерпретацию коэффициенту полной регрессии, сделать прогноз на 5 лет вперед.
7. Построить модель машинного обучения по методу «Случайный лес», оценить точность классификации разными метриками.

Примерный перечень вопросов к зачету

1. Понятие открытых данных.
2. Источники данных и их виды.
3. Качество данных.
4. Достоверность источников данных.
5. Правовое регулирование данных.
6. Открытые данные в России.
7. Открытые данные за рубежом.
8. Основные статистики данных.
9. Ряды распределения и их виды.
10. Ранжированные ряд распределения, когда используется, алгоритм построения.
11. Дискретный ряд распределения, сфера применения, алгоритм построения.
12. Интервальный ряд распределения, применение и алгоритм построения.
13. Показатели центральной тенденции.
14. Среднее значение. Расчет в ранжированном, дискретном и интервальном ряду распределения.
15. Модальное значение. Расчет в ранжированном, дискретном и интервальном ряду распределения.
16. Медианное значение. Расчет в ранжированном, дискретном и интервальном ряду распределения.
17. Показатели вариации.

18. Размах вариации, среднее линейное отклонение, дисперсия, среднее квадратическое отклонение. Интерпретация и алгоритм расчетов в дискретном ряду распределения.
19. Коэффициент осцилляции, относительно линейное отклонение, коэффициент вариации. Интерпретация и алгоритм расчетов в дискретном ряду распределения.
20. Размах вариации, среднее линейное отклонение, дисперсия, среднее квадратическое отклонение. Интерпретация и алгоритм расчетов в интервальном ряду распределения.
21. Коэффициент осцилляции, относительно линейное отклонение, коэффициент вариации. Интерпретация и алгоритм расчетов в интервальном ряду распределения.

Пример практических работ

Практическая работа № 1. Основные термины и понятия открытых данных.

Вопросы для обсуждения:

1. Понятие открытых данных.
2. Источники данных и их виды.
3. Качество данных.
4. Достоверность источников данных.
5. Правовое регулирование данных.
6. Открытые данные в России.
7. Открытые данные за рубежом.

Практическая работа № 2. Построение рядов распределения в Excel, Python, R

1. На основе исходных данных, выбрать показатели, подходящие для построения ранжированного, дискретного и интервального рядов распределения.
2. Написать программу, позволяющую автоматизировать процесс построения ранжированного, дискретного и интервального рядов распределения на Python или R.
3. Проверить правильность построения в Excel.

Вопросы к защите:

1. Основные статистики данных.
2. Ряды распределения и их виды.

Практическая работа № 3. Описательные статистики: средняя, мода, медиана в Excel, Python, R

1. На основе результатов выполнения предыдущего задания, написать программу, позволяющую автоматизировать процесс расчетов показателей центральной тенденции для ранжированного, дискретного и интервального рядов распределения на Python или R.
2. Проверить правильность построения в Excel.

Вопросы к защите:

1. Показатели центральной тенденции.
2. Среднее значение. Расчет в ранжированном, дискретном и интервальном ряду распределения.
3. Модальное значение. Расчет в ранжированном, дискретном и интервальном ряду распределения.
4. Медианное значение. Расчет в ранжированном, дискретном и интервальном ряду распределения.

Практическая работа № 4. Показатели вариации в Excel, Python, R

1. На основе результатов выполнения предыдущих заданий, написать программу, позволяющую автоматизировать процесс расчетов показателей вариации для ранжированного, дискретного и интервального рядов распределения на Python или R.
2. Проверить правильность построения в Excel.

Вопросы к защите:

1. Показатели вариации.
2. Размах вариации, среднее линейное отклонение, дисперсия, среднее квадратическое отклонение. Интерпретация и алгоритм расчетов в дискретном ряду распределения.
3. Коэффициент осцилляции, относительно линейное отклонение, коэффициент вариации. Интерпретация и алгоритм расчетов в дискретном ряду распределения.
4. Размах вариации, среднее линейное отклонение, дисперсия, среднее квадратическое отклонение. Интерпретация и алгоритм расчетов в интервальном ряду распределения.
5. Коэффициент осцилляции, относительно линейное отклонение, коэффициент вариации. Интерпретация и алгоритм расчетов в интервальном ряду распределения.

Практическая работа № 5. Загрузка данных в Python, работа с типом Series

1. Загрузите spyder. Загрузите файл с исходными данными Вашего варианта в формате xlsx, csv и txt. При загрузке автоматически присвойте первой строке названия столбцам.
2. Обратитесь только к 3 строке таблицы.
3. Обратитесь к 2 и 54 значениям таблицы.
4. Вывести первые 7 значений.
5. Вывести последние 12 значений.
6. Вместо поменять местами четные значения.
7. Добавить на каждое нечетное значение, данные по соседнему варианту.
8. Рассчитать среднюю, моду, медиану, дать интерпретацию

Вопросы к защите:

1. Обращения к столбцам таблицы
2. Обращение к строкам таблицы
3. Добавление новых строк.

Практическая работа № 6. Загрузка данных в Python, работа с Data Frame

1. Загрузите spyder. Загрузите файл с исходными данными Вашего варианта.
2. Провести аналитику всех показателей (минимум, максимум, среднее, мода, медиана).
3. Проверить на соответствие нормальному закону распределения.

Вопросы к защите:

1. Способы загрузки данных в среду
2. Библиотеки для работы с таблицами

Практическая работа № 7. Загрузка данных в R.

4. Загрузите spyder. Загрузите файл с исходными данными Вашего варианта.
5. Провести аналитику всех показателей (минимум, максимум, среднее, мода, медиана).
6. Проверить на соответствие нормальному закону распределения.

Вопросы к защите:

3. Способы загрузки данных в среду
4. Библиотеки для работы с таблицами

Практическая работа № 8. Обработка данных в Data Frame

1. На основе практической работы 6 провести анализ исходных данных, а именно:
2. Рассчитывать показатели размеров, специализации, интенсификации предприятий, сделать выводы

Вопросы к защите:

1. Способы загрузки данных в среду
2. Библиотеки для работы с таблицами

Практическая работа № 9. Библиотеки для визуализации данных в Python

1. По данным практической работы 6 подстроить линейный график, точечную дигамму, столбчатую диаграмму, круговую диаграмму, гистограмму.
2. Изменить настройки отображения диаграмм.

Вопросы к защите:

1. Графические возможности Python
2. Редактирование графиков
3. Добавление нескольких графиков
4. Составные графики
5. Модуль matplotlib

Практическая работа № 10. Библиотеки для визуализации данных в R

1. По данным практической работы 6 подстроить линейный график, точечную дигамму, столбчатую диаграмму, круговую диаграмму, гистограмму.
2. Изменить настройки отображения диаграмм.

Вопросы к защите:

1. Графические возможности R

2. Редактирование графиков
3. Добавление нескольких графиков
4. Составные графики
5. Модуль ggplot, ggplot2 и др.

Практическая работа № 11. Построение моделей корреляционно-регрессионного анализа

1. По данным работы 6 построить парные и множественные модели регрессии.
2. Оценить значимость моделей и параметров.
3. Отобрать качественные модели, дать интерпретацию коэффициенту полной регрессии.
4. Сделать прогноз.

Вопросы к защите:

1. Оценка качества уравнения регрессии
2. Оценка значимости уравнения в целом
3. Оценка значимости параметров уравнения регрессии
4. Отбор факторов в уравнение регрессии
5. Предпосылки МНК

Практическая работа № 12. Классификация методом машинного обучения «Дерево решений»

1. Загрузить исходные данные.
2. Провести обучение модели методом «Дерево решений» на полном массиве данных.
3. На основе обученной модели провести классификацию данных.
4. Оценить точность классификации различными метриками.

Вопросы к защите:

1. Виды деревьев решений
2. Интерпретация результатов
3. Сокращение числа «ветвей»
4. Оценка качество построенного дерева
5. Особенности бинарных деревьев решений

Практическая работа № 13. Другие методы машинного обучения.

1. На основе данных задачи 12 провести обучение моделей другими методами.
2. Оценить точность классификации моделей.
3. Сравнить результаты с методом «Дерево решений».

Вопросы к защите:

1. Категориальные переменные
2. Особенности выбора зависимой переменной
3. Оценка качества модели
4. Интерпретация результатов
5. Применение как метод классификации

Практическая работа № 14. Оптимизация моделей, подбор параметров.

1. На основе данных задачи 12 провести оптимизацию моделей «Дерево решений», путем подбора «лучшего» критерия разбиения и глубины дерева.
2. Сравнить результаты с задачами 12 и 13.

Вопросы к защите:

1. Оптимальная модель
2. Критерий Джини
3. Критерий Энтропии
4. Глубина дерева

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

Текущий контроль знаний, умений и навыков проводится в форме контрольной работа 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.

Максимальная оценка за выполнение каждой работы составляет 10 баллов. Задержка выполнения индивидуального практического задания на одну неделю штрафует одним баллом. Критерии оценки выполненных работ приведены в таблице:

Оценка	Критерии оценивания
«10» (отлично)	оценку «отлично» заслуживает студент, выполнивший практическую работу полностью, в работе корректно применены статистические методы, нет ошибок в расчетах, сделаны глубокие выводы. Студент дал полные ответы на все заданные вопросы по работе. Недостатков по оформлению работы не имеется.
«8-9» (хорошо)	оценку «хорошо» заслуживает студент, выполнивший практическую работу полностью, имеются недочеты в применении статистических методов, проведенном анализе и полученных выводах. Студент дал верные ответы на все заданные вопросы по работе. Недостатков по оформлению работы не имеется.
«6-7» (удовлетворительно)	оценку «удовлетворительно» заслуживает студент, выполнивший практическую работу полностью, но имеются существенные недочеты в применении отдельных статистических методов и полученных выводах. Студент дал верные ответы не на все вопросы. По оформлению работы имеются недостатки.
«1-5» (неудовлетворительно)	оценку «неудовлетворительно» заслуживает студент, выполнивший практическую работу не по своему варианту или с грубейшими нарушениями применения методов и последовательности анализа.

Контрольная работа оценивается по 100-бальной системе.

Оценка «отлично» выставляется студенту, если удельный вес правильных ответов составляет более 84%

Оценка «хорошо» выставляется студенту, если удельный вес правильных ответов составляет 75 – 84%

Оценка «удовлетворительно» выставляется студенту, если удельный вес правильных ответов составляет 60 – 74%

Оценка «неудовлетворительно» выставляется студенту, если удельный вес правильных ответов составляет более 0 – 59%

Итоговая оценка учитывает результаты рейтинговой системы контроля знаний (вклад 80%), результаты контрольных работ по всем темам (вклад 20%)
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: 24.09.2022). — 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: 24.09.2022). — Режим доступа: для авториз. пользователей.
3. Полупанов, Д. В. Программирование в Python 3 : учебное пособие / Д. В. Полупанов, С. Р. Абдюшева, А. М. Ефимов. — Уфа : БашГУ, 2020. — 164 с. — ISBN 978-5-7477-5230-6. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/179915> (дата обращения: 24.09.2022). — Режим доступа: для авториз. пользователей.
4. Фролов, Ю. В. Управление знаниями : учебник для вузов / Ю. В. Фролов. — 2-е изд., испр. и доп. — Москва : Издательство Юрайт, 2022. — 324 с. — (Высшее образование). — ISBN 978-5-534-05521-4. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/493448> (дата обращения: 24.09.2022).

7.2 Дополнительная литература

1. Боровков, А. А. Математическая статистика : учебник для вузов / А. А. Боровков. — 5-е изд., стер. — Санкт-Петербург : Лань, 2021. — 704 с. — ISBN 978-5-8114-7677-0. — Текст : электронный // Лань : электронно-

- библиотечная система. — URL: <https://e.lanbook.com/book/164711> (дата обращения: 24.09.2022). — Режим доступа: для авториз. пользователей.
2. Кремер, Н. Ш. Эконометрика : учебник и практикум для вузов / Н. Ш. Кремер, Б. А. Путко ; под редакцией Н. Ш. Кремера. — 4-е изд., испр. и доп. — Москва : Издательство Юрайт, 2022. — 308 с. — (Высшее образование). — ISBN 978-5-534-08710-9. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/488678> (дата обращения: 24.09.2022).
 3. Нестеров, С. А. Основы интеллектуального анализа данных. Лабораторный практикум : учебное пособие / С. А. Нестеров. — Санкт-Петербург : Лань, 2020. — 40 с. — ISBN 978-5-8114-4509-7. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/130181> (дата обращения: 24.09.2022). — Режим доступа: для авториз. пользователей.

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

1. Язык программирования Python. URL: <https://www.python.org/>(открытый доступ)
2. Анаконда. URL: <https://www.anaconda.com/distribution/>(открытый доступ)
3. Язык программирования R. URL: <https://www.r-project.org/>(открытый доступ)
4. RStudio: <https://posit.co/download/rstudio-desktop/?azure-portal=true>(открытый доступ)
5. MS Office: <https://www.microsoft.com/ru-ru/microsoft-365/try?market=ru>
6. Официальный сайт Росстата. URL: <http://www.gks.ru/> (открытый доступ)
7. Официальный сайт Центрального Банка России. URL: <http://www.cbr.ru> (открытый доступ)

9. Перечень программного обеспечения

Таблица 9

Перечень программного обеспечения

№ п/п	Наименование раздела учебной дисциплины	Наименование программы	Тип программы	Автор	Год разработки
1	Раздел 1. Введение в науку о данных. Основы работы в Python Раздел 2. Основные статистики в Python Раздел 3. Графические возможности Python Раздел 4. Статистические методы в Python Раздел 5. Машинное обучение в Python	Anaconda, RStudio, MS Office	расчётная	Anaconda Enterprise, Posit, PBC, Microsoft	2019, 2011, 2016

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)	Оснащенность специальных помещений и помещений для самостоятельной работы
1	2
<i>учебная аудитория для проведения занятий лекционного типа, учебная аудитория для групповых и индивидуальных консультаций, учебная аудитория для текущего контроля и промежуточной аттестации (2й учебный корпус, 102 ауд.)</i>	<ol style="list-style-type: none"> 1. Экран с электроприводом 1 шт. (Инв. №558771/2) 2. Проектор 1 шт. (без инв. №) – приобретался не за счет средств вуза 3. Вандалоустойчивый шкаф 1 шт. (Инв.№558850/7) 4. Системный блок iP-4 541 3200 Mhz/1024 Mb/ 80 Gb / DVD-R с монитором 1 шт. (Инв. №558777/9) 5. Стенд «Сергеев Сергей Степанович 1910-1999» 1 шт. (Инв.№591013/25) 6. Огнетушитель порошковый 1 шт. (Инв. №559527) 7. Подвесное крепление к огнетушителю 1 шт. (Инв. № 559528) 8. Жалюзи 2шт. (Инв. №1107-221225, Инв. №1107-221225) 9. Лавка 20 шт. 10. Стол аудиторный 20 шт. 11. Стол для преподавателя 1 шт. 12. Стул 2 шт. 13. Доска маркерная 1 шт. 14. Трибуна напольная 1 шт. (без инв. №)
<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 студентам по освоению дисциплины

Приступая к изучению дисциплины «Технологии работы с открытыми данными», студенты должны ознакомиться с учебной программой, учебной, научной и методической 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.

Предполагается, что студент выполняет практическое задание в аудитории, дома оформляет и готовится по теоретическим вопросам к защите работы на следующем занятии.

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

Комплексное освоение студентами учебной дисциплины «Технологии работы с открытыми данными» предполагает изучение рекомендуемой учебно-методической литературы, подготовку к практическим занятиям, самостоятельную работу при выполнении практических заданий, домашних заданий.

На первом занятии преподаватель закрепляет за каждым студентом номер варианта для выполнения индивидуальных работ (как правило, номер варианта соответствует порядковому номеру студента в журнале преподавателя). По каждой индивидуальной работе должна быть поставлена оценка по факту ее защиты. Защиту рекомендуется проводить на следующем после получения задания занятии. Преподаватель обязан проверить соответствие выполненного задания исходным данным варианта студента. Таким образом, исключается вероятность плагиата.

В рамках курса предусмотрены формы работы студентами в малых группах, ориентированные на развитие навыков взаимодействия у студентов при решении профессиональных задач.

Преподаватель должен стимулировать студентов к занятию научно-исследовательской работой, изучению научной эконометрической литературы, в т.ч. отечественной и зарубежной периодики.

Студент может провести собственное статистическое наблюдение за социально-экономическими явлениями, представляющими его научный интерес, построить статистическую модель, сделать прогноз. В случае надлежащего качества, его работа может быть заслушана на научном кружке кафедры или на студенческой научной конференции. По решению кафедры, студенты, занявшие призовые места на научных студенческих конференциях, могут освободиться от сдачи зачета по дисциплине.

Программу разработал (и):

Харитонов А.Е., к.э.н., доцент

Ульянкин А.Е., ассистент



РЕЦЕНЗИЯ

на рабочую программу дисциплины Б1.В.25 «Технологии работы с открытыми данными» ОПОП ВО по направлению 09.03.03 «Прикладная информатика», направленность «Прикладная информатика в экономике» (квалификация выпускника – бакалавр)

Коломеевой Еленой Сергеевной, доцентом кафедры финансов ФГБОУ ВО г. Москвы «РГАУ-МСХА имени К.А. Тимирязева», кандидатом экономических наук (далее по тексту рецензент), проведено рецензирование рабочей программы дисциплины «Технологии работы с открытыми данными» ОПОП ВО по направлению 09.03.03 «Прикладная информатика», направленность «Прикладная информатика в экономике» (уровень обучения - бакалавриат) разработанной в ФГБОУ ВО «Российский государственный аграрный университет – МСХА имени К.А. Тимирязева», на кафедре статистики и кибернетики (разработчики – Харитоновна Анна Евгеньевна, кандидат экономических наук, доцент кафедры статистики и кибернетики и Ульянов Александр Евгеньевич, ассистент кафедры статистики и кибернетики).

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

1. Предъявленная рабочая программа дисциплины «Технологии работы с открытыми данными» (далее по тексту Программа) соответствует требованиям ФГОС ВО по направлению 09.03.03 «Прикладная информатика». Программа содержит все основные разделы, соответствует требованиям к нормативно-методическим документам.

2. Представленная в Программе актуальность учебной дисциплины в рамках реализации ОПОП ВО не подлежит сомнению – дисциплина относится к дисциплинам части, формируемой участниками образовательных отношений учебного цикла учебного цикла — Б1.В.

3. Представленные в Программе цели дисциплины соответствуют требованиям ФГОС ВО направления 09.03.03 «Прикладная информатика».

4. В соответствии с Программой за дисциплиной «Технологии работы с открытыми данными» закреплено 2 компетенции (4 индикаторов). Дисциплина «Технологии работы с открытыми данными» и представленная Программа способна реализовать их в объявленных требованиях.

5. Результаты обучения, представленные в Программе в категориях знать, уметь, владеть соответствуют специфике и содержанию дисциплины и демонстрируют возможность получения заявленных результатов.

6. Общая трудоёмкость дисциплины «Технологии работы с открытыми данными» составляет 3 зачётные единицы (108 часов).

7. Информация о взаимосвязи изучаемых дисциплин и вопросам исключения дублирования в содержании дисциплин соответствует действительности. Дисциплина «Технологии работы с открытыми данными» взаимосвязана с другими дисциплинами ОПОП ВО и Учебного плана по направлению 09.03.03 «Прикладная информатика» и возможность дублирования в содержании отсутствует.

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

9. Программа дисциплины «Технологии работы с открытыми данными» предполагает проведение занятий в интерактивной форме.

10. Виды, содержание и трудоёмкость самостоятельной работы студентов, представленные в Программе, соответствуют требованиям к подготовке выпускников, содержащимся во ФГОС ВО направления 09.03.03 «Прикладная информатика».

11. Представленные и описанные в Программе формы текущей оценки знаний (защита работ), соответствуют специфике дисциплины и требованиям к выпускникам.

Форма промежуточного контроля знаний студентов, предусмотренная Программой, осуществляется в форме зачета в 4 семестре, что *соответствует* статусу дисциплины, как дисциплины части, формируемой участниками образовательных отношений учебного цикла – Б1.В ФГОС ВО направления 09.03.03 «Прикладная информатика».

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

13. Учебно-методическое обеспечение дисциплины представлено: основной литературой – 4 источника (базовый учебник), дополнительной литературой – 3 наименования Интернет-ресурсы – 6 источников и соответствует требованиям ФГОС ВО направления 09.03.03 «Прикладная информатика».

14. Материально-техническое обеспечение дисциплины соответствует специфике дисциплины «**Технологии работы с открытыми данными**» и обеспечивает использование современных образовательных, в том числе интерактивных методов обучения.

15. Методические рекомендации студентам и методические рекомендации преподавателям по организации обучения по дисциплине дают представление о специфике обучения по дисциплине «**Технологии работы с открытыми данными**».

ОБЩИЕ ВЫВОДЫ

На основании проведенного рецензирования можно сделать заключение, что характер, структура и содержание рабочей программы дисциплины «**Технологии работы с открытыми данными**» ОПОП ВО по направлению 09.03.03 «Прикладная информатика», направленность «**Прикладная информатика в экономике**» (квалификация выпускника – бакалавр), разработанная Харитоновой А. Е., к.э.н., доцентом кафедры статистики и кибернетики и Ульяновским А.Е. ассистентом 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

(подпись)



«24» августа 2022 г.