

Документ подписан простой электронной подписью.
Информация о владельце:

ФИО: Кандрашина Елена Александровна

Должность: И.о. ректора ФГАОУ ВО «Самарский государственный экономический университет»

Дата подписания: 07.08.2025 11:55:43

Уникальный программный ключ:

2db64eb9605ce27edd3b8e8fdd32c70e0674ddd2

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

Институт Институт менеджмента
Кафедра Прикладной информатики

УТВЕРЖДЕНО

Ученым советом Университета
(протокол № 10 от 22 мая 2025 г.)

РАБОЧАЯ ПРОГРАММА

Наименование дисциплины	Б1.В.ДЭ.04.01 Большие данные
Основная профессиональная образовательная программа	38.03.05 Бизнес-информатика программа Бизнес-информатика

Квалификация (степень) выпускника бакалавр

Самара 2025

Содержание (рабочая программа)

Стр.

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

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

1. Место дисциплины в структуре ОП

Дисциплина Большие данные входит в часть, формируемая участниками образовательных отношений (дисциплина по выбору) блока Б1. Дисциплины (модули)

Предшествующие дисциплины по связям компетенций: Основы алгоритмизации и программирования

Последующие дисциплины по связям компетенций: Тестирование информационных систем, Тестирование при разработке ПО, Веб-разработка, Машинное обучение и анализ больших данных, Проектирование цифровых сервисов в управлении

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

Изучение дисциплины Большие данные в образовательной программе направлено на формирование у обучающихся следующих компетенций:

Профессиональные компетенции (ПК):

ПК-3 - Способен разрабатывать прототипы ИС и интегрировать ИС с существующими ИС заказчика в рамках регламента выполнения работ по созданию (модификации) ИС

Планируемые результаты обучения по программе	Планируемые результаты обучения по дисциплине		
ПК-3	ПК-3.1: Знать:	ПК-3.2: Уметь:	ПК-3.3: Владеть (иметь навыки):
	устройство и функционирование современных ИС, программные средства и платформы инфраструктуры информационных технологий организаций -партнеров, кодирование на языках программирования в рамках выполнения работ по созданию (модификации) и сопровождению ИС, инструменты и методы моделирования бизнес-процессов ИС	разрабатывать прототипы ИС и интегрировать ИС с существующими ИС заказчика в рамках регламента выполнения работ по созданию (модификации) ИС	навыками моделирования бизнес-процессов заказчика ИС в типовой ИС, навыками проведения анализа функциональных разрывов и формулирование предложения заказчику ИС по изменению его бизнес-процессов для реализации их автоматизации в типовой ИС

3. Объем и виды учебной работы

Учебным планом предусматриваются следующие виды учебной работы по дисциплине:

Очная форма обучения

Виды учебной работы	Всего час/ з.е.
	Сем 5
Контактная работа, в том числе:	36.15/1
Занятия лекционного типа	18/0.5
Занятия семинарского типа	18/0.5
Индивидуальная контактная работа (ИКР)	0.15/0

Самостоятельная работа:	53.85/1.5
Промежуточная аттестация	18/0.5
Вид промежуточной аттестации: Зачет	Зач
Общая трудоемкость (объем части образовательной программы): Часы	108
Зачетные единицы	3

4. Содержание дисциплины

4.1. Разделы, темы дисциплины и виды занятий:

Тематический план дисциплины Большие данные представлен в таблице.

Разделы, темы дисциплины и виды занятий Очная форма обучения

№ п/п	Наименование темы (раздела) дисциплины	Контактная работа				Самостоятельная работа	Планируемые результаты обучения в соотношении с результатами обучения по образовательной программе
		Лекции	Занятия семинарского типа	ИКР	ГКР		
			Практич. занятия				
1.	Методика проведения комплексного многомерного сравнительного анализа (МСА). Подготовка и первичная обработка данных.	9	18			23	ПК-3.1, ПК-3.2, ПК-3.3
2.	Современные инструменты и алгоритмы анализа BIG DATA.	9	18			30,85	ПК-3.1, ПК-3.2, ПК-3.3
	Контроль	18					
	Итого	18	18	0.15		53.85	

4.2 Содержание разделов и тем

4.2.1 Контактная работа

Тематика занятий лекционного типа

№п/п	Наименование темы (раздела) дисциплины	Вид занятия лекционного типа*	Тематика занятия лекционного типа
1.	Методика проведения комплексного многомерного сравнительного анализа (МСА). Подготовка и первичная обработка данных.	лекция	Основные понятия BIG DATA. Технологии и инструментарий анализа больших данных.
		лекция	Методика проведения комплексного многомерного сравнительного анализа (МСА). Применение статистических методов для технологий анализа больших данных. Общая схема проведения комплексного многомерного статистического анализа.
		лекция	Система показателей - основа повышения качества и точности моделей, основанных на анализе больших данных. Подготовка датасета.
		лекция	Python. Библиотеки для работы с данными. Методика и особенности

			проведения разведочного анализа с помощью Python. Визуализация данных и результатов как инструмент.
2.	Современные инструменты и алгоритмы анализа BIG DATA.	лекция	Методы многомерного сравнительного анализа и нормализация данных. Задача построения рейтинга.
		лекция	Методы многомерной классификации. Иерархический кластерный анализ и итеративные методы. Реализация в Python: методика обучения без учителя.
		лекция	Построение регрессионных моделей с помощью Python.
		лекция	Loginom - платформа для анализа данных. Введение, основы работы.
		лекция	Loginom. практические примеры анализа экономических показателей.

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

Тематика занятий семинарского типа

№п/п	Наименование темы (раздела) дисциплины	Вид занятия семинарского типа**	Тематика занятия семинарского типа
1.	Методика проведения комплексного многомерного сравнительного анализа (МСА). Подготовка и первичная обработка данных.	практическое занятие	Лабораторная работа 1. Разработка системы показателей. Сбор и подготовка данных.
		практическое занятие	Лабораторная работа 1. Разработка системы показателей. Сбор и подготовка данных.
		практическое занятие	Лабораторная работа 1. Язык Python. Работа в Анаконде. Формирование датасета. Работа с записями датафрейма.
		практическое занятие	Лабораторная работа 2. Разведочный анализ данных. Язык Python. Описательные статистики. Интерпретация результатов.
2.	Современные инструменты и алгоритмы анализа BIG DATA.	практическое занятие	Лабораторная работа 2. Разведочный анализ данных. Язык Python. Описательные статистики. Интерпретация результатов.
		практическое занятие	Лабораторная работа 2. Разведочный анализ данных. Язык Python. Проверка на однородность, коэффициент вариации. Определение выбросов.
		практическое занятие	Лабораторная работа 2. Разведочный анализ данных. Язык Python. Проверка на однородность, коэффициент вариации. Определение выбросов.
		практическое занятие	Лабораторная работа 2. Проверка совокупностей на соответствие нормальному закону распределения. Гистограммы и графики распределения в Python.
		практическое занятие	Лабораторная работа 2. Выявление взаимосвязей, коэффициент

			корреляции. Построение диаграмм рассеяния и визуализация в Python. Интерпретация результатов.
--	--	--	---

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

Иная контактная работа

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

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

4.2.2 Самостоятельная работа

№п/п	Наименование темы (раздела) дисциплины	Вид самостоятельной работы ***
1.	Методика проведения комплексного многомерного сравнительного анализа (МСА). Подготовка и первичная обработка данных.	- подготовка доклада - подготовка электронной презентации - тестирование
2.	Современные инструменты и алгоритмы анализа BIG DATA.	- подготовка доклада - подготовка электронной презентации - тестирование

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

5. Материально-техническое и учебно-методическое обеспечение дисциплины

5.1 Литература:

Основная литература

1. Парфенов, Ю. П. Постреляционные хранилища данных : учебное пособие для вузов / Ю. П. Парфенов ; под научной редакцией Н. В. Папуловской. — Москва : Издательство Юрайт, 2025. — 97 с. — (Высшее образование). — ISBN 978-5-534-21173-3. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/559502>

Дополнительная литература

1. Гордеев, С. И. Организация баз данных : учебник для вузов / С. И. Гордеев, В. Н. Волошина. — 2-е изд., испр. и доп. — Москва : Издательство Юрайт, 2025. — 691 с. — (Высшее образование). — ISBN 978-5-534-21115-3. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/559377>

2. Гордеев, С. И. Организация баз данных : учебник для вузов / С. И. Гордеев, В. Н. Волошина. — 2-е изд., испр. и доп. — Москва : Издательство Юрайт, 2025. — 691 с. — (Высшее образование). — ISBN 978-5-534-21115-3. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/559377>

Литература для самостоятельного изучения

1.

5.2. Перечень лицензионного программного обеспечения

1. Astra Linux Special Edition «Смоленск», «Орел»; РедОС ; ОС "Альт Рабочая станция" 10; ОС "Альт Образование" 10
2. МойОфис Стандартный 2, МойОфис Образование, Р7-Офис Профессиональный, МойОфис Стандартный 3, МойОфис Профессиональный 3
3. Anaconda Navigator, Jupyter Notebook, Loginom

5.3 Современные профессиональные базы данных, к которым обеспечивается доступ обучающихся

1. Профессиональная база данных «Информационные системы Министерства экономического развития Российской Федерации в сети Интернет» (Портал «Официальная Россия» - <http://www.gov.ru/>)

2. Государственная система правовой информации «Официальный интернет-портал правовой информации» (<http://pravo.gov.ru/>)

3. Профессиональная база данных «Финансово-экономические показатели Российской Федерации» (Официальный сайт Министерства финансов РФ - <https://www.minfin.ru/ru/>)

4. Профессиональная база данных «Официальная статистика» (Официальный сайт Федеральной службы государственной статистики - <http://www.gks.ru/>)

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

1. Справочно-правовая система «Консультант Плюс»

2. Справочно-правовая система «ГАРАНТ-Максимум»

5.5. Специальные помещения

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

5.6 Лаборатории и лабораторное оборудование

Лаборатория информационных технологий в профессиональной деятельности	Комплекты ученической мебели Мультимедийный проектор Доска Экран
---	---

	Компьютеры с выходом в сеть «Интернет» и ЭИОС СГЭУ Лабораторное оборудование
--	--

6. Фонд оценочных средств по дисциплине Большие данные:

6.1. Контрольные мероприятия по дисциплине

Вид контроля	Форма контроля	Отметить нужное знаком « + »
Текущий контроль	Оценка докладов	+
	Тестирование	+
	Лабораторные работы	+
Промежуточный контроль	Зачет	+

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

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

Профессиональные компетенции (ПК):

ПК-3 - Способен разрабатывать прототипы ИС и интегрировать ИС с существующими ИС заказчика в рамках регламента выполнения работ по созданию (модификации) ИС

Планируемые результаты обучения по программе	Планируемые результаты обучения по дисциплине		
	ПК-3.1: Знать:	ПК-3.2: Уметь:	ПК-3.3: Владеть (иметь навыки):
	устройство и функционирование современных ИС, программные средства и платформы инфраструктуры информационных технологий организаций -партнеров, кодирование на языках программирования в рамках выполнения работ по созданию (модификации) и сопровождению ИС, инструменты и методы моделирования бизнес-процессов ИС	разрабатывать прототипы ИС и интегрировать ИС с существующими ИС заказчика в рамках регламента выполнения работ по созданию (модификации) ИС	навыками моделирования бизнес-процессов заказчика ИС в типовой ИС, навыками проведения анализа функциональных разрывов и формулирование предложения заказчику ИС по изменению его бизнес-процессов для реализации их автоматизации в типовой ИС
Пороговый	устройство и функционирование	собирать и анализировать информацию	навыки анализа информации.

	современных ИС, программные средства и платформы инфраструктуры информационных технологий организаций -партнеров		моделирования бизнес-процессов заказчика
Стандартный (в дополнение к пороговому)	кодирование на языках программирования в рамках выполнения работ по созданию (модификации) и сопровождению ИС	использовать современные информационные технологии и программные средства;	навыками работы с программными средствами для обработки данных в соответствии с поставленной задачей
Повышенный (в дополнение к пороговому, стандартному)	инструменты и методы моделирования бизнес-процессов ИС	интерпретации и анализа информации, отбирать технологии работы с информацией в зависимости от класса задач в области данных анализировать информацию, результаты расчетов, обосновывать полученные выводы	навыками проведения анализа функциональных разрывов и формулирование предложения заказчику ИС по изменению его бизнес-процессов для реализации их автоматизации в типовой ИС

6.3. Паспорт оценочных материалов

№ п/п	Наименование темы (раздела) дисциплины	Контролируемые планируемые результаты обучения в соотношении с результатами обучения по программе	Вид контроля/используемые оценочные средства	
			Текущий	Промежуточный
1.	Методика проведения комплексного многомерного сравнительного анализа (МСА). Подготовка и первичная обработка данных.	ПК-3.1, ПК-3.2, ПК-3.3	Тестирование Лабораторные работы	Зачет
2.	Современные инструменты и алгоритмы анализа BIG DATA.	ПК-3.1, ПК-3.2, ПК-3.3	Тестирование Лабораторные работы	Зачет

6.4.Оценочные материалы для текущего контроля

Примерная тематика докладов

Раздел дисциплины	Темы
Методика проведения комплексного многомерного сравнительного анализа (МСА). Подготовка и первичная обработка	1. Big data(Большие данные) 2. Международные проекты в области Big Data 3. Big Data/ Основы Hadoop. Базовый набор компонентов Hadoop. 4. Big Data: обработка больших объемов данных 5. Большие данные (Big Data). Назначение, применение, перспектива. 6. Применение big data в маркетинговых исследованиях потребителей.

данных.	Большие данные - Big Data - Статкомитет СНГ
Современные инструменты и алгоритмы анализа BIG DATA.	1. Статистический анализ данных в связях с общественностью 2. RMarkdown, R и ggplot 3. Прогнозные модели посещаемости сайта в RStudio 4. RStudio, можно будет построить интерактивную диаграмму 5. Data Science для карьериста 6. Среда разработки RStudio 7. Особенности эконом. состояния регионов страны. Статистическая оценка эконом. состояния регионов страны в RStudio. 8. Исследование однородности основных финансовых показателей регионов РФ 9. Оценка качества жизни в субъектах РФ на основе анализа путем многомерной кластеризации статистических данных 10. Сравнительный анализ регионов Киргизии по уровню экономического развития за 2021, 2022г. 11. Сравнительный анализ уровня жизни по странам 12. Дифференциация регионов Российской Федерации по уровням и факторам прибыли финансовых организаций 13. Анализ уровня жизни по параметрам трудовой сферы районов Турции 14. Экономическое исследование регионов Казахстана 15. Исследование социально-экономических показателей Армении в аналитической платформе Loginom 16. Разведочный анализ развития строительства РФ

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

<https://lms2.sseu.ru/course/index.php?categoryid=1909>

Система статистических показателей – это...

+ совокупность взаимосвязанных между собой показателей, представляющих собой целостную и логичную систему

набор невязанных показателей, характеризующих количественную меру по каждому показателю

Показатели в системе статистических показателей делятся на группы:

+ факторные

+ результативные

интегральные

дискретные

Не является основным требованием к системе статистических показателей ...:

репрезентативность данных

достоверность данных

нормирование данных

+ классификация данных

Нормировать данные по показателю, означает:

+ привести данные к одинаковым единицам измерения

представить данные в числовом виде

представить оперативные данные

Нормированными показателями являются:

+ общий строительный объём введенных зданий, м³ на душу населения
общий строительный объём введенных зданий, тыс. м³
+ среднемесячная заработная плата, руб.
Численность постоянного населения, тыс. чел.

Метод многомерной средней заключается в ...
в выявлении приоритетности регионов по совокупности определенных позиций
+ в замене разномасштабных абсолютных значений данных нормализованными, вычисленными
как отношение значения к среднему значению по показателю

К методам многомерного сравнительного анализа относятся:
+ метод Паттерн
+ метод суммы мест
метод наименьших квадратов

Не является методом многомерного сравнительного анализа:
метод Паттерн
метод суммы мест
+ метод наименьших квадратов

Мода - это
+ наиболее часто встречающееся значение переменной
среднее по порядку значение
сумма всех значений отнесенная, к общему числу наблюдений

медиана – это ...
наиболее часто встречающееся значение переменной
+ среднее по порядку значение
сумма всех значений отнесенная, к общему числу наблюдений

коэффициент вариации – это
+ отношение стандартного отклонения к среднему арифметическому, выраженному в процентах
стандартное отклонение, деленное на квадратный корень выборки
разность между наибольшим и наименьшим значениями набора данных

коэффициент вариации, рассчитанный по показателю «среднемесячная заработная плата, руб.»
равен 97%. Это означает, что совокупность:
однородная
+ неоднородная

коэффициент вариации, рассчитанный по показателю «среднемесячная заработная плата, руб.»
равен 27%. Это означает, что совокупность:
+ однородная
неоднородная

коэффициент вариации, рассчитанный по показателю «Доля занятого населения» равен 45%. Это

означает, что совокупность:

однородная

+ неоднородная

коэффициент вариации, рассчитанный по показателю «Доля занятого населения» равен 31%. Это означает, что совокупность:

+однородная

неоднородная

к этапам разведочного анализа относятся:

+ исследование степени однородности статистических совокупностей

+ выявление «выбросов» и исключение их из совокупности данных

+ анализ на соответствие эмпирических распределений нормальному закону

Корреляционно-регрессионный анализ

Факторный анализ

этапом разведочного анализа не является:

исследование степени однородности статистических совокупностей

выявление «выбросов» и исключение их из совокупности данных

анализ на соответствие эмпирических распределений нормальному закону

+ Корреляционно-регрессионный анализ

Критерий Пирсона используется:

+ для проверки гипотез о независимости или соответствии распределений

проверки значимости уравнения регрессии

для проверки тесноты связи между функцией и несколькими факторами-аргументами

нормальное распределение это:

+ это непрерывное распределение вероятностей, где значения симметрично сгруппированы вокруг среднего, а вероятность отклонений уменьшается по мере удаления от него

все значения в заданном диапазоне имеют одинаковую вероятность; график плотности

вероятности представляет собой прямую линию

моделирует время между событиями в процессе, происходящем с постоянной средней частотой

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

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

-<p>math.sin</p>

-<p>m.sin</p>

-<p>(m)sin</p>

-<p>sin</p>

}

В соответствии с рекомендациями Code Style переменные в Python принято называть

::В соответствии с рекомендациями Code Style переменные в Python принято называть::[html]<p>В соответствии с рекомендациями Code Style переменные в Python принято называть</p>{

-<p>в PascalStyle</p>

-<p>в camelStyle</p>

-<p>в стиле snake_case</p>

-<p>в стиле UPPER_SNAKE_CASE</p>
</p>

В чём состоит свойство расширяемости записей СУБД?

::В чём состоит свойство расширяемости записей СУБД?::[html]<p>В чём состоит свойство расширяемости записей СУБД?</p>{

-<p>СУБД не имеет чёткой структуры, поэтому любую запись можно расширить</p>
-<p>повышение отказоустойчивости системы при добавлении новых записей в СУБД</p>
-<p>в любую таблицу СУБД можно добавить новую колонку, предварительно изменив структуру этой таблицы</p>
-<p>СУБД имеет чёткую, но расширяемую структуру, в каждую запись можно добавить новую колонку, также как и узнать значение любой записи по добавленной колонке</p>
</p>

Величина погрешности характеристики при использовании метода Монте-Карло

::Величина погрешности характеристики при использовании метода

Монте-Карло::[html]<p>Величина погрешности характеристики при использовании метода Монте-Карло</p>{

-<p>обратно пропорциональна числу испытаний</p>
-<p>прямо пропорциональна числу испытаний</p>
-<p>прямо пропорциональна корню из числа испытаний</p>
-<p>обратно пропорциональна корню из числа испытаний</p>
</p>

Выберите корректную запись цикла для перебора значений от 0 до 10 в Python.

::Выберите корректную запись цикла для перебора значений от 0 до 10 в

Python.::[html]<p>Выберите корректную запись цикла для перебора значений от 0 до 10 в Python.</p>{

-<p>for i \: \ - 0 to 10 do</p>
-<p>for (int i \ - 0; i < \ - 10; i++)</p>
-<p>for i in range(11)\:</p>
-<p>for i in (0, 10)\:</p>
</p>

Выберите цикл, который будет бесконечным.

::Выберите цикл, который будет бесконечным.::[html]<p>Выберите цикл, который будет бесконечным.</p>{

-<p>for i in range(-1)\:</p>
-<p>while True\:</p>
-<p>while False\:</p>
</p>

Дистрибутив, содержащий основные библиотеки Python для работы с данными и научных исследований...

::Дистрибутив, содержащий основные библиотеки Python для работы с данными и научных исследований....::[html]<p>Дистрибутив, содержащий основные библиотеки Python для работы с данными и научных исследований, называется</p>{

-<p>Anaconda</p>
-<p>SciPython</p>
-<p>Jupyter Notebook</p>
-<p>SciLab</p>
</p>

Доверительная вероятность характеризует:

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

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

Доверительный интервал случайной величины есть величина...

```
::Доверительный интервал случайной величины есть величина...::[html]<p>Доверительный
интервал случайной величины есть величина...</p>{
  -<p>случайная</p>
  -<p>детерминированная</p>
  -<p>случайная или детерминированная, в зависимости от объема исходной выборки</p>
}
```

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

```
::Известно, что между величинами X и Y существует положительная связь. В каких пределах
находится...::[html]<p>Известно, что между величинами X и Y существует положительная связь. В
каких пределах находится парный коэффициент корреляции?</p>{
  -<p>от -1 до 0</p>
  -<p>от 0 до 1</p>
  -<p>от -1 до 1</p>
}
```

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

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

Имеется файловая переменная my_file. Команда my_file.closed позволяет узнать...

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

Имеется файловая переменная my_file. Команда my_file.name позволяет узнать...

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

Имеется файловая переменная my_file. Команда my_file.read() позволяет...

```
::Имеется файловая переменная my_file. Команда my_file.read() позволяет...::[html]<p>Имеется
файловая переменная my_file. Команда my_file.read() позволяет...</p>{
```

-<p>открыть файл для чтения</p>
-<p>прочитать из файла данные до конца строки</p>
-<p>прочитать из файла один байт данных</p>
-<p>прочитать весь файл до конца</p>

}

Имена переменных могут включать:

::Имена переменных могут включать\:::[html]<p>Имена переменных могут включать\:</p>{

-<p>русские буквы</p>
-<p>латинские буквы</p>
-<p>пробелы</p>
-<p>скобки, знаки + \ - ! ? b др.</p>
-%25%<p>цифры</p>
-%25%<p>знак подчеркивания</p>

}

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

::Инвестиционный фонд интересуется тем, почему часть финансируемых им проектов успешно переходят....:[html]<p>Инвестиционный фонд интересуется тем, почему часть финансируемых им проектов успешно переходят на второй год, а часть — нет. К какому типу относится эта задача анализа данных?</p>{

-<p>построение решающего правила</p>
-<p>классификация</p>
-<p> поиск информативных признаков</p>
-<p>цензурирование</p>

}

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

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

-<p>номинальная</p>
-<p>абсолютная</p>
-<p>порядковая</p>
-<p>бинарная</p>

}

К слабоструктурированным относятся задачи, которые содержат

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

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

}

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

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

-<p>метод прогнозирования</p>
-<p>метод моделирования</p>
-<p>метод оптимизации</p>
-<p>метод алгоритмизации</p>

-<p>метод деривации</p>
}

Как называется нарушение допущения о независимости остатков?

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

-<p>Мультиколлинеарность</p>
-<p>Автокорреляция</p>
-<p>Гетероскедастичность</p>
-<p>Гомоскедастичность</p>
}

Как называется нарушение допущения о постоянстве дисперсии остатков?

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

-<p>Мультиколлинеарность</p>
-<p>Автокорреляция</p>
-<p>Гетероскедастичность</p>
-<p>Гомоскедастичность</p>
}

Как правильно присвоить p значение типа данных, имеющих две координаты x

-1 p;и p;y -...

::Как правильно присвоить p значение типа данных, имеющих две координаты x\

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

-<p>p(x,y)<-c(1,2)</p>
-<p> p \ - c(y\ -2, x\ -1)</p>
-<p> p \ - c(x\ -1, y\ -2)</p>
-<p>p.x\ -1.y\ -2</p>
}

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

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

-<p>Hbase</p>
-<p>Cassandra</p>
-p> Redis</p>
-p> BigTable</p>

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

Раздел дисциплины	Задачи
Методика проведения комплексного многомерного сравнительного анализа (МСА). Подготовка и первичная обработка данных.	ТАА 1 <i>Лабораторная работа №1 Формирование датасета (2 часа)</i> 1. На основании логико-содержательного анализа отобрать показатели и сформировать систему в соответствии с заданным вариантом в MS Excel. 2. Сформируйте файл датасет в MS Excel, который: • содержит результативные (выход) и факторные (вход) показатели, изучаемого процесса (пример сформированной базы и требования к её наполнению смотрите в описании);

	• количество и название показателей определяется студентом, исходя из принципов целесообразности, необходимости и достаточности; • показатели берутся по регионам РФ за одинаковый период, например, за 2021 год; • берутся только регионы, без данных по федеральным округам • в базу данных следует брать нормированные показатели. 3. Дать посмотреть систему показателей преподавателю, для их конечного согласования. 4. В Python создать файл с информационной системой показателей <i>Лабораторная работа №2 Разведочный анализ данных (4 часа)</i> Использовать базу данных, сформированную в процессе выполнения лабораторной работы №1. Задания выполняются в Python 1. Определите описательные статистики. 2. Интерпретируйте результаты (описательные статистики) о степени однородности по одному какому-либо выбранному вами показателю 3. Построить гистограммы по каждому показателю. Выявить соответствие эмпирических данных теоретическому закону распределения. Сделать вывод о доле нормально или «близконормально» распределенных совокупностей показателей. 4. Определите коэффициент вариации для каждого показателя. Сделать выводы об однородности совокупностей исследуемой системы показателей. 5. Провести анализ данных на точки «выбросов» для неоднородных совокупностей. 6. Получить матричные диаграммы рассеивания. Выявить форму зависимости одной переменной от другой (если они существуют).
Современные инструменты и алгоритмы анализа BIG DATA.	ТАА 2 <i>Лабораторная работа № 3 «Разработка системы определения рейтинговых оценок экономических объектов непараметрическими методами многомерного сравнительного анализа»</i> <i>Цель:</i> Освоить непараметрические методы агрегирования информации, методы «сжатия» информационного пространства в среде Excel. 1. Использовать БД, сформированную в среде Excel в лабораторной работе №1. 2. На последующих листах необходимо реализовать непараметрические методы: ранжирования, итерационный метод многомерной средней, метод «Паттерн» (листы

именовать названиями методов).

3. Определить интегральные оценки и ранги регионов каждым методом.

4. Построить итоговую таблицу с помощью ссылок на ячейки, содержащую ранги, полученные каждым методом. Сравнить результаты с помощью коэффициента Спирмена (для расчета коэффициентов завести новую таблицу), рассчитанного по формуле:

$$r_{sp} = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)},$$

где d_i^2 – разности рангов, полученных двумя сравниваемыми методами;

n – число пар рангов (число регионов)

Результаты расчетов оформить в виде квадратной матрицы парных коэффициентов ранговых корреляций.

Лабораторная работа №4 Кластерный анализ и получение устойчивых типологических групп

Использовать базу данных, сформированную в процессе выполнения лабораторной работы №1. Задания выполняются в Python.

1. Используя иерархическую кластеризацию, построить дендрограмму («Лекция 4_2_методы обучения без учителя.docx», стр 12 ...)
 1. Определить уровень выделения кластеров.
 2. Определить состав кластеров.
 3. Определить число кластеров (для п.3).
2. Используя метод К-средних, получить разбиение на кластеры, число кластеров задаёте, по данным п.2.3.
3. Определить состав кластеров.
4. Определить средние значения показателей по кластерам на графике (постройте его). Идентифицировать кластеры по интегральным оценкам.
5. Сравнить результаты, полученные методами интегральной средней и k-средних. Выявить устойчивые типологические группы.

ТААЗ

Loginom. Задача 1.

1. создайте пакет в **Loginom** и импортируйте файл **Товары.txt**;
2. исключите из набора записи, в которых для товара отсутствует артикул;
3. с помощью функций компонента **Калькулятор** рассчитайте, сколько месяцев прошло от даты последней продажи каждого товара до 01.04.2018;
4. отсортируйте набор данных по количеству месяцев по убыванию;
5. Добавьте в набор поле логического типа Вывод из продажи и установите значение true для товаров, у которых до 01.04.2018 прошло более 10 месяцев.

	Loginom. Задача 2. Работа в Калькуляторе и подключение библиотек Рассмотрим основы работы на следующем примере: - создайте в Loginom новый пакет и импортируйте файл Вероятности визитов.txt; - рассчитайте количество клиентов, у которых в идентификаторе Клиент.Код присутствуют два и более нулей; !! для решения данной задачи не требуется создать сложный сценарий, достаточно внимательно изучить функции узла Калькулятор; - подключите к вашему пакету библиотеку Loginom Silver Kit; - отберите всех клиентов, которые входят в тестовое множество; - далее используйте внешний компонент AUC из библиотеки Loginom Silver Kit, вам потребуются следующие настройки в мэппинге полей: Отклик – Событие; Вероятность визита – Оценка - после настройки входного порта достаточно запустить узел на выполнение. Loginom. Задача 3. Сценарий, узел и основы работы с модулями Loginom Рассмотрим основы работы со сценариями на следующем примере: - создайте в Loginom новый пакет и импортируйте файл Вероятности визитов.txt; - Разработайте сценарий, который рассчитывает долгосрочную стоимость клиента CLV – Customer Lifetime Value по следующей формуле: $CLV = m \left(\frac{r}{1 + i - r} \right),$ где: - m – ожидаемый годовой доход с одного клиента в рублях; - r- коэффициент удержания клиента, от 0 до 1; - i – годовая ставка дисконтирования в долях от 0 до 1; - рассчитайте для каждого клиента из файла Вероятности визитов.txt значение показателя CLV, если ставка дисконтирования равна 12%, а коэффициент удержания и ожидаемый годовой доход зависят от типа клиента: - r=0,9 и m=7200 – для клиентов с типом 1; - r=0,7 и m=9500 – для клиентов с типом 2; - отсортируйте по убыванию CLV; - проделайте то же самое, рассчитав CLV по другой формуле, с учетом ежегодного роста прибыли на g%: $CLV = m \left(\frac{r}{1 + i - r(1 + g)} \right)$ - при расчетах возьмите g=8%.
--	---

6.5. Оценочные материалы для промежуточной аттестации

Фонд вопросов для проведения промежуточного контроля в форме зачета

Раздел дисциплины	Вопросы
Методика проведения комплексного многомерного сравнительного анализа (МСА).	- В чем принципиальное отличие концепции Big Data от традиционного подхода BI? - Понятие явной (выраженной) и скрытой (структурной) информации. - Определение контент-анализа. - Каковы основные понятия контент-анализа?

	5. Какие существуют виды контент-анализа? 6. Какие существуют этапы контент-анализа? 7. Каковы основные признаки, характеризующие «Большие данные»? 8. Модели развертывания облачных хранилищ. 9. Модели обслуживания облачных хранилищ. 10. Постановка и описание проблемы «последней мили». 11. Безопасность, производительность и надежность при работе с облачными данными. 12. Экономическая составляющая облачных подходов. 13. Способы машинного обучения. 14. Основные фазы обработки «больших данных» 15. Понятие системы показателей. Основные требования к формированию информационной базы. 16. Что такое разведочный анализ? 17. Какие этапы проведения разведочного анализа? 18. Для чего используются описательные статистики? 19. Что такое нормальное распределение? 20. Что такое критерий согласия Пирсона? Для чего он используется? 21. Что такое χ^2? 22. Какая совокупность является однородной? 23. Для чего используется коэффициент вариации? 24. Как рассчитывается коэффициент вариации? 25. Что означает понятие «выброс»? 26. Как определяется «выброс»?
Современные инструменты и алгоритмы анализа BIG DATA.	27. Методы МСА. Общая схема проведения комплексного многомерного сравнительного анализа. 28. Методы МСА. Предпосылки и этапы реализации МСА. 29. Методы многомерного сравнительного анализа (суммы мест, многомерной средней, паттерн, ранжирования). 30. Итерационный метод многомерного сравнительного анализа. 31. Понятие кластерного анализа. Основное назначение методов кластерного анализа. 32. Кластерный анализ. Типы методов многомерной классификации. 33. Понятие меры сходства двух объектов. Выбор метрики или меры близости объектов. 34. Методы объединения объектов в кластеры и присоединения объекта к кластеру 35. Общий алгоритм классификации методами иерархического кластерного анализа 36. Дивизимные методы многомерной кластеризации 37. Особенности итеративных методов кластерного анализа. Общие сведения о методе К – средних. 38. Алгоритм многомерной кластеризации объектов методом К-средних. 39. Критерии качества классификации методов кластерного анализа. 40. Разведочный анализ как этап формирования однородной совокупности. Выявление степени однородности совокупности, точек «выбросов», степени соответствия распределения эмпирических данных теоретическим законам распределения. 41. Исследование эмпирических распределений совокупностей. 42. Вычисление основных статистик. Получение графических результатов. Формулировка выводов о существовании или отсутствии закономерностей в исследуемых процессах. 43. Загрузка встроенных в пакеты датасетов 44. Импорт данных из файла формата csv 45. Импорт данных из файла формата MS Excel 46. Оценивание линейной регрессии 47. Прогнозирование по оценённой модели линейной регрессии 48. Тестирование гипотез о коэффициент 49. F-тест для коэффициентов регрессии 50. Диагностические тесты

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

Шкала и критерии оценивания

Оценка	Критерии оценивания для мероприятий контроля с применением 2-х балльной системы
«зачтено»	ПК-3
«не зачтено»	Результаты обучения не сформированы на пороговом уровне