

Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Гриб Владислав Валерьевич
Должность: Ректор
Дата подписания: 03.11.2023 16:59:02
Уникальный программный ключ:
637517d24e103c3db032acf37e06498e1c5bb2f5ab80c39ebfad7f47095447



Образовательное частное учреждение высшего образования
«МОСКОВСКИЙ УНИВЕРСИТЕТ ИМЕНИ А.С. ГРИБОЕДОВА»
(ИМПЭ им. А.С. Грибоедова)

Институт международной экономики, лидерства и менеджмента

УТВЕРЖДАЮ
Директор института
международной экономики,
лидерства и менеджмента
_____ А.А. Панарин
«28» сентября 2023 г.

Рабочая программа дисциплины
СТАТИСТИЧЕСКИЕ МЕТОДЫ АНАЛИЗА ДАННЫХ

Направление подготовки 09.03.03 Прикладная информатика
(уровень бакалавриат)

Направленность (профиль):
«Анализ данных»

Форма обучения: очная

Москва

Рабочая программа дисциплины «Статистические методы анализа данных». Направление подготовки 09.03.03 Прикладная информатика, направленность (профиль): «Анализ данных» / А.А. Шестемиров – М.: ИМПЭ им. А.С. Грибоедова. – 46 с.

Рабочая программа дисциплины составлена на основании федерального государственного образовательного стандарта высшего образования – бакалавриат по направлению подготовки 09.03.03 Прикладная информатика, утвержденного приказом Министерства образования и науки Российской Федерации от «19» сентября 2017 № 922 (с изменениями и дополнениями) и Профессионального стандарта «Программист», утвержденного приказом Министерства труда и социальной защиты Российской Федерации от 18 ноября 2013 г. № 679н (зарегистрирован Министерством юстиции Российской Федерации 18 декабря 2013 г., регистрационный № 30635), с изменением, внесенным приказом Министерства труда и социальной защиты Российской Федерации от 12 декабря 2016 г. № 727н (зарегистрирован Министерством юстиции Российской Федерации 13 января 2017 г., регистрационный № 45230), Профессионального стандарта «Специалист по информационным системам», утвержденный приказом Министерства труда и социальной защиты Российской Федерации от 18 ноября 2014 г. № 896н (зарегистрирован Министерством юстиции Российской Федерации 24 декабря 2014 г., регистрационный № 35361), с изменением, внесенным приказом Министерства труда и социальной защиты Российской Федерации от 12 декабря 2016 г. № 727н (зарегистрирован Министерством юстиции Российской Федерации 13 января 2017 г., регистрационный № 45230)

Разработчики:

К.э.н. А.А. Шестемиров

Ответственный рецензент:

Назарова Н.А., к.э.н., доцент, заместитель руководителя департамента налогов и налогового администрирования Финансового университета при Правительстве Российской Федерации

(Ф.И.О., уч. степень, уч. звание, должность)

Рабочая программа дисциплины рассмотрена и одобрена на заседании кафедры инновационного менеджмента и предпринимательства 15.09.2023г., протокол №2

Заведующий кафедрой _____ / _____ /к.э.н. А.А. Шестемиров/

(подпись)

Согласовано от Библиотеки _____ /О.Е. Степкина/

(подпись)

РАЗДЕЛ 1. ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Целью освоения дисциплины «Статистические методы анализа данных» является изучение современной методологии анализа данных и формирование навыков статистического анализа данных с целью разработки алгоритмов математического обеспечения информационных систем.

Задачи, решаемые в ходе освоения программы дисциплины:

- ознакомить обучающихся с методами обработки данных, со сферами применения статистических методов анализа данных в информационных системах;
- сформировать навыки решения профессионально-ориентированных задач с применением статистических методов обработки и анализа данных.

РАЗДЕЛ 2. ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ, СООТНЕСЕННЫЕ С ПЛАНИРУЕМЫМИ РЕЗУЛЬТАТАМИ ОСВОЕНИЯ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ

Категория (группа) компетенций	Код компетенции	Формулировка компетенции	Индикаторы достижения компетенции (для планирования результатов обучения по элементам образовательной программы и соответствующих оценочных средств)
Универсальные компетенции			
Системное и критическое мышление	УК-1	Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач	ИУК-1.1 Знать методы поиска информации для решения поставленной задачи ИУК-1.2 Уметь проводить критический анализ и синтез информации ИУК-1.3 Владеть системным подходом для решения поставленных задач

ПРОФЕССИОНАЛЬНЫЕ КОМПЕТЕНЦИИ

Код, наименование профессиональных компетенций	Трудовые функции (код, наименование)/уровень (подуровень) квалификации	Индикаторы достижения компетенции (для планирования результатов обучения по элементам образовательной программы и соответствующих оценочных средств)
ОТФ (код, наименование) / Профессиональный стандарт (код, наименование)		
Тип(ы) задач(и) профессиональной деятельности		
D. Разработка требований и проектирование программного обеспечения/ 06.001 Программист		
Проектный		
ПК-1. Способность проектировать прикладное программное обеспечение	D/03.6 Проектирование программного обеспечения	ИПК-1.1 Знать принципы построения архитектуры программного обеспечения и виды архитектуры программного обеспечения; типовые решения, библиотеки программных модулей, шаблоны, классы объектов, используемые при разработке программного обеспечения; методы и средства проектирования программного обеспечения; методы и средства проектирования баз данных; методы и средства проектирования программных интерфейсов ИПК-1.2 Уметь использовать существующие типовые решения и шаблоны проектирования программного обеспечения; применять методы и средства проектирования программного обеспечения, структур данных, баз данных, программных интерфейсов; осуществлять коммуникации с заинтересованными сторонами ИПК-1.3 Владеть разработкой, изменением и согласованием архитектуры программного обеспечения с системным аналитиком и архитектором программного обеспечения; системой проектирования структур данных; системой проектирования баз данных; системой проектирования программных интерфейсов; оценкой и согласованием сроков выполнения поставленных задач
С Выполнение работ и управление работами по созданию (модификации) и сопровождению ИС, автома-		

<p>тизирующих задачи организационного управления и бизнес-процессы / 06.015 Профессиональный стандарт «Специалист по информационным системам»</p> <p>производственно-технологический , научно-исследовательский</p>		
<p>ПК-2. Способность осуществлять инженерно-технологическую поддержку планирования управления требованиями по созданию (модификации) и сопровождению ИС</p>	<p>С/10.6 Инженерно-технологическая поддержка планирования управления требованиями</p>	<p>ИПК-2.1. Знать инструменты и методы управления требованиями; предметной областью автоматизации; возможностями ИС; источниками информации, необходимой для профессиональной деятельности; современный отечественный и зарубежный опыт в профессиональной деятельности; инструменты и методы моделирования бизнес-процессов организации; основы реинжиниринга бизнес-процессов организации; управление содержанием проекта: документирование требований, систему анализа продукта, модерлируемых совещаний; управление качеством: контрольные списки, верификация, валидация (приемо-сдаточные испытания); управление коммуникациями в проекте: базовые навыки управления (в том числе проведение презентаций, проведение переговоров, публичные выступления); культуру речи; правила деловой переписки.</p> <p>ИПК-2.2. Умеет анализировать входные данные; планировать работы.</p> <p>ИПК-2.3. Владеть системным подходом для выбора технологии управления требованиями; представлениями исходных данных для разработки плана управления требованиями; системой согласования в части инженерно-технологического обеспечения плана управления требованиями с заинтересованными сторонами</p>
<p>ПК-3. Способность применять системный подход для составления технико-экономического обоснования проектных решений и технического задания на разработку информационной системы</p>	<p>С/01.6 Определение первоначальных требований заказчика к ИС и возможности их реализации в ИС на этапе предконтрактных работ;</p>	<p>ИПК3.1. Знает возможности типовой ИС; предметную область автоматизации; методы выявления требований; технологии межличностной и групповой коммуникации в деловом взаимодействии, основы конфликтологии; технологии подготовки и проведения презентаций; архитектуру, устройство и функционирование вычислительных систем; коммуникационное оборудование; сетевые протоколы; основы современных операционных систем; основы современных систем управления базами данных; устройство и функционирование современных ИС; современные стандарты информационного взаимодействия систем; программные средства и платформы инфраструктуры информационных технологий организаций; современные подходы и стандарты автоматизации организации (например, CRM, MRP, ERP..., ITIL, ITSM); основы теории систем и системного анализа; методики описания и моделирования бизнес-процессов, средства моделирования бизнес-процессов; системы классификации и кодирования информации, в том числе присвоение кодов документам и элементам справочников; отраслевая нормативная техническая документация; источники информации, необходимой для профессиональной деятельности; современный отечественный и зарубежный опыт в профессиональной деятельности; формирование и механизмы рыночных процессов организации; основы менеджмента, в том числе менеджмента качества; основы бухгалтерского учета и отчетности организаций; основы налогового законодательства Российской Федерации; основы управленческого учета; основы финансового учета и бюджетирования; основы международных стандартов финансовой отчетности (МСФО); основы управления торговлей, поставками и запасами; основы организации производства; основы управления взаимоотношениями с клиентами и заказчиками (CRM); основы управления персоналом, включая вопросы оплаты труда; основы теории управления; современные инструменты и методы управления организацией, в том числе методы плани-</p>

		<p>рования деятельности, распределения поручений, контроля исполнения, принятия решений; методология ведения документооборота в организациях; инструменты и методы определения финансовых и производственных показателей деятельности организаций; основы организационной диагностики; инструменты и методы моделирования бизнес-процессов организации; основы реинжиниринга бизнес-процессов организации; управление содержанием проекта документирование требований, анализ продукта, модерлируемые совещания; культура речи; правила деловой переписки.</p> <p>ИПК-3.2 Умеет проводить переговоры; проводить презентации; подготавливать протоколы мероприятий.</p> <p>ИПК 3.3. Владеет основами выявления первоначальных требований заказчика к ИС; системой информирования заказчика о возможностях типовой ИС и вариантах ее модификации; определениями возможностей достижения соответствия ИС первоначальным требованиям заказчика; основами составления протоколов переговоров с заказчиком</p>
--	--	--

РАЗДЕЛ 3. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ

Дисциплина «Статистические методы анализа данных» изучается в пятом, шестом семестрах, относится к Б1.В. Части учебного плана, формируемой участниками образовательных отношений, Блока Б.1 «Дисциплины (модули)».

Общая трудоемкость дисциплины составляет 10 з.е.

Знания, умения, навыки, опыт практической деятельности, приобретенные при освоении настоящей дисциплины, необходимы для успешного освоения следующих дисциплин: «Математическая теория риска», «Системы анализа данных», «Избранные вопросы анализа данных».

РАЗДЕЛ 4. ОБЪЕМ (ТРУДОЕМКОСТЬ) ДИСЦИПЛИНЫ

(ОБЩАЯ, ПО ВИДАМ УЧЕБНОЙ РАБОТЫ, ВИДАМ ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ)

Трудоемкость дисциплины и виды учебной работы (очная форма обучения)

З.е.	Всего часов	Контактная работа			Часы СР на подготовку кур.раб.	Иная СР	Контроль
		Занятия лекционного типа	Занятия семинарского типа				
			Лабораторные	Практические/Семинарские			
5 семестр							
5	180	24	40			114	2 Зачет с оценкой
6 семестр							
				-			
Всего по дисциплине							
5	180	24	40			80	36 Экзамен

СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

№ п/п	Наименование разделов и тем дисциплины	Содержание темы
5 семестр		
Раздел №1 «Введение в статистические методы анализа данных»		

№ п/п	Наименование разделов и тем дисциплины	Содержание темы
1	Тема 1.1 Введение в статистические методы анализа данных.	<p>Изучаемые вопросы:</p> <ol style="list-style-type: none"> 1. Предмет дисциплины «Статистические методы анализа данных». 2. Основные типы данных. 3. Основные этапы эмпирического исследования. <p>Вопросы для самостоятельного изучения:</p> <ol style="list-style-type: none"> 1. Информационный подход к анализу данных.
2	Тема 1.2 Методы первичной обработки и представления данных.	<p>Изучаемые вопросы:</p> <ol style="list-style-type: none"> 1. Методы предобработки данных. 2. Группировка данных. 3. Методы нормализации или стандартизации данных. <p>Вопросы для самостоятельного изучения:</p> <ol style="list-style-type: none"> 1. Приложение методов обработки данных к решению практических задач.
3	Тема 1.3 Методы выявления и устранения аномальных наблюдений.	<p>Изучаемые вопросы:</p> <ol style="list-style-type: none"> 1. Аномальные наблюдения или «выбросы». 2. Классификация аномальных наблюдений. 3. Критерий Смирнова-Граббса. 4. Критерий Титьена-Мура. 5. Критерий Ирвина. <p>Вопросы для самостоятельного изучения:</p> <ol style="list-style-type: none"> 1. Графические методы выявления аномальных наблюдений. 2. Методы устранения аномальных наблюдений
Раздел №2 «Множественный корреляционно-регрессионный анализ. Классическая модель множественной регрессии»		
4	Тема 2.1 Корреляционно-регрессионный анализ. Модель множественной регрессии.	<p>Изучаемые вопросы:</p> <ol style="list-style-type: none"> 1. Общий вид модели множественной регрессии. 2. Классификация регрессии и корреляции. 3. Модель множественной линейной регрессии. Матричная форма. 4. МНК-оценки теоретических коэффициентов. <p>Вопросы для самостоятельного изучения:</p> <ol style="list-style-type: none"> 1. Этапы построения модели множественной регрессии.
5	Тема 2.2 Фиктивные переменные. Критерий Чоу проверки однородности.	<p>Тема: Фиктивные переменные. Критерий Чоу проверки однородности.</p> <p>Изучаемые вопросы:</p> <ol style="list-style-type: none"> 1. Фиктивные переменные во множественной регрессии. 2. Критерий Чоу анализа структурных изменений. <p>Вопросы для самостоятельного изучения:</p> <ol style="list-style-type: none"> 1. Методы формирования выборки наблюдений фиктивных переменных для построения регрессии.
6	Тема 2.3 Множественные коэффициенты корреляции.	<p>Изучаемые вопросы:</p> <ol style="list-style-type: none"> 1. Частный коэффициент корреляции, его свойства. 2. Множественный коэффициент корреляции. Свойства. 3. Частный коэффициент детерминации. <p>Вопросы для самостоятельного изучения:</p> <ol style="list-style-type: none"> 1. Коэффициент детерминации. Свойства.
7	Тема 2.4 Статистический анализ адекватности модели множественной регрессии.	<p>Изучаемые вопросы:</p> <ol style="list-style-type: none"> 1. Теорема Гаусса-Маркова. 2. Статистические свойства МНК-оценок. 3. Проверка гипотез о статистической значимости теоретических коэффициентов регрессии. 4. Проверка гипотезы о статистической значимости уравнения регрессии. 5. Дисперсионный анализ. 6. Дисперсионное отношение Фишера <p>Вопросы для самостоятельного изучения:</p> <ol style="list-style-type: none"> 1. Доверительные интервалы для теоретических коэффициентов регрессии.

№ п/п	Наименование разделов и тем дисциплины	Содержание темы
		2. Практическая реализация проверки адекватности модели множественной регрессии
8	Тема 2.5 Мультиколлинеарность факторов.	Изучаемые вопросы: 1. Мультиколлинеарность факторов. 2. Признаки мультиколлинеарности. 3. Методы выявления мультиколлинеарности. 4. Методы устранения мультиколлинеарности. Вопросы для самостоятельного изучения: 1. Последствия мультиколлинеарности.
Раздел №3 «Обобщенная линейная модель множественной регрессии. Регрессионные модели с автокоррелированными остатками»		
9	Тема 3.1 Обобщенная линейная модель множественной регрессии. Обобщенный МНК	Изучаемые вопросы: 1. Обобщенная линейная модель множественной регрессии. 2. Обобщенный МНК. Теорема Айткена. Вопросы для самостоятельного изучения: 1. Взвешенный МНК.
10	Тема 3.2 Гетероскедастичность. Методы выявления гетероскедастичности.	Изучаемые вопросы: 1. Гетероскедастичность. Графический анализ остатков на гетероскедастичность. 2. Тест Гольфельда-Квандта. 3. Тест Уайта. Вопросы для самостоятельного изучения: 1. Тест Глейзера.
11	Тема 3.3 Взвешенный МНК. Устранение гетероскедастичности.	Изучаемые вопросы: 1. Двухуровневая дисперсия. Взвешенный МНК. 2. Пропорциональность дисперсии случайных ошибок значениям объясняющей переменной. 3. Зависимость дисперсии случайных ошибок от значений объясняющей переменной. Вопросы для самостоятельного изучения: 1. Примеры устранения гетероскедастичности.
12	Тема 3.4 Регрессионные модели с автокоррелированными остатками.	Изучаемые вопросы: 1. Понятие автокорреляции остатков. Причины автокорреляции остатков. 2. Графический анализ остатков. 3. Критерий Дарбина-Уотсона. 4. Тест Бреуша-Годфри. Вопросы для самостоятельного изучения: 1. Примеры практической реализации критериев проверки автокорреляции остатков.
13	Тема 3.5 Устранение автокорреляции	Изучаемые вопросы: 1. Метод взвешенного МНК для случая модели с автокоррелированными остатками. 2. Методы построения оценки коэффициента автокорреляции первого порядка. Вопросы для самостоятельного изучения: 1. Методы построения оценки коэффициента автокорреляции первого порядка.
6 семестр		
Раздел №4 «Кластерный анализ»		
14	Тема 4.1 Кластерный анализ. Основные этапы кластерного анализа.	Изучаемые вопросы: 1. Понятие кластерного анализа. 2. Основные этапы кластерного анализа. 3. Представление данных в задаче кластеризации. Вопросы для самостоятельного изучения:

№ п/п	Наименование разделов и тем дисциплины	Содержание темы
		1. Основные задачи и приложения кластерного анализа.
15	Тема 4.2 Метрики расстояний. Свойства кластеров. Методы кластерного анализа.	Изучаемые вопросы: 1. Метрики расстояний. Свойства метрик. 2. Свойства кластеров. Вопросы для самостоятельного изучения: 1. Классификация методов кластерного анализа.
16	Тема 4.3 Расстояния между кластерами. Методы группировки кластеров.	Изучаемые вопросы: 1. Расстояния между кластерами. 2. Методы группировки кластеров. Вопросы для самостоятельного изучения: 1. Дисперсионные методы группировки кластеров.
17	Тема 4.4 Иерархические методы кластеризации. Агломеративные методы.	Изучаемые вопросы: 1. Иерархические методы кластеризации. 2. Алгоритм реализации агломеративных методов. Вопросы для самостоятельного изучения: 1. Иерархические агломеративные методы кластеризации. Практическое приложение.
18	Тема 4.5 Иерархические дивизимные методы кластеризации.	Изучаемые вопросы: 1. Иерархические дивизимные методы. 2. Алгоритм реализации дивизимных методов кластеризации. Вопросы для самостоятельного изучения: 1. Иерархические дивизимные методы кластеризации. Практическое приложение.
19	Тема 4.6 Примеры реализации иерархических методов кластеризации.	Изучаемые вопросы: 1. Пример реализации агломеративного алгоритма. 2. Пример реализации дивизимного алгоритма. Вопросы для самостоятельного изучения: 1. Дендограмма.
20	Тема 4.7 Итеративные методы. Метод k-средних.	Изучаемые вопросы: 1. Итеративные методы. 2. Метод k-средних. Вопросы для самостоятельного изучения: 1. Пример реализации метода k-средних.
21	Тема 4.8 Метод k-средних. Критерии качества кластеризации.	Изучаемые вопросы: 1. Функционалы качества разбиения. 2. Критерии качества кластеризации. Вопросы для самостоятельного изучения: 1. Статистические критерии качества разбиения.
22	Тема 4.9 Метод нечетких k-средних	Изучаемые вопросы: 1. Метод нечетких k-средних. 2. Функционалы качества разбиения. Вопросы для самостоятельного изучения: 1. Пример реализации метода нечетких k-средних.
Раздел №5 «Методы классификации»		
23	Тема 5.1 Понятие методов классификации.	Изучаемые вопросы: 1. Постановка задачи классификации. 2. Основные методы классификации. Вопросы для самостоятельного изучения: 1. Примеры задач классификации.
24	Тема 5.2 Логистическая регрессия.	Изучаемые вопросы: 1. Общий вид логистической регрессии. 2. Оценка логистической регрессии. 3. Понятие сбалансированной выборки. Вопросы для самостоятельного изучения: 1. Методы оценки параметров логистической регрессии.

№ п/п	Наименование разделов и тем дисциплины	Содержание темы
25	Тема 5.3 Метрики качества классификации.	Изучаемые вопросы: 1. Метрики качества классификации. 2. Матрица ошибок. 3. Доля правильных ответов, точность, полнота. Вопросы для самостоятельного изучения: 1. Примеры расчета метрик качества классификации.
26	Тема 5.4 Метрики качества классификации. Регуляризация в задаче регрессии.	Изучаемые вопросы: 1. ROC-кривая, площадь под кривой. 2. L1-регуляризация. 3. L2-регуляризация. Вопросы для самостоятельного изучения: 1. Построение ROC-кривой.
27	Тема 5.5 Методы деревьев решений.	Изучаемые вопросы: 1. Деревья решений. 2. Основные алгоритмы построения деревьев решений. Вопросы для самостоятельного изучения: 1. Основные алгоритмы построения деревьев решений.
28	Тема 5.6 Ансамбли деревьев решений: Случайный лес.	Изучаемые вопросы: 1. Понятие ансамблей решений. 2. Случайный лес. Вопросы для самостоятельного изучения: 1. Пример практической реализации алгоритма случайного леса.
29	Тема 5.7 Ансамбли деревьев решений: Градиентный бустинг.	Изучаемые вопросы: 1. Градиентный бустинг. 2. Реализация ансамблей решений. Вопросы для самостоятельного изучения: 1. Пример практической реализации алгоритма градиентного бустинга.

Перечень разделов (модулей), тем дисциплины и распределение учебного времени по разделам/темам дисциплины, видам учебных занятий (в т.ч. контактной работы), видам текущего контроля очная форма обучения

Разделы / Темы	Контактная работа			Часы СР на подготовку кур.р.	Иная СР	Контроль	Всего часов
	Занятия лекционного типа	Занятия семинарского типа					
		Лаб.р	Прак. /сем.				
5 семестр							
Раздел 1 «Введение в статистические методы анализа данных»	6	8			38		52
Тема 1.1 Введение в статистические методы анализа данных.	2				13		15
Тема 1.2 Методы первичной обработки и представления данных.	2	4			13		19
Тема 1.3 Методы выявления и устранения аномальных наблюдений.	2	4			12		18
Раздел №2 «Множественный корреляционно-регрессионный анализ. Классическая модель мно-	10	20			38		68

Разделы / Темы	Контактная работа			Часы СР на под- готов- ку кур. р.	Иная СР	Кон- троль	Всего часов	
	Заня- тия лек- цион- ного типа	Занятия се- минарского типа						Кон- такт- ная работа по кур.р
		Лаб. р	Прак · /сем.					
жественной регрессии»								
Тема 2.1 Корреляционно-регрессионный анализ. Модель множественной регрессии.	2	4			8		14	
Тема 2.2 Фиктивные переменные. Критерий Чоу проверки однородности.	2	4			8		14	
Тема 2.3 Множественные коэффициенты корреляции.	2	4			8		14	
Тема 2.4 Статистический анализ адекватности модели множественной регрессии.	2	4			7		13	
Тема 2.5 Мультиколлинеарность факторов.	2	4			7		13	
Раздел №3 «Обобщенная линейная модель множественной регрессии. Регрессионные модели с автокоррелированными остатками»	8	12			38		58	
Тема 3.1 Обобщенная линейная модель множественной регрессии. Обобщенный МНК	1				8		9	
Тема 3.2 Гетероскедастичность. Методы выявления гетероскедастичности.	1				8		9	
Тема 3.3 Взвешенный МНК. Устранение гетероскедастичности.	2	4			8		14	
Тема 3.4 Регрессионные модели с автокоррелированными остатками.	2	4			7		13	
Тема 3.5 Устранение автокорреляции	2	4			7		13	
Зачет с оценкой						2	2	
Итого за 5 семестр	24	40			114	2	180	
6 семестр								
Раздел №4 «Кластерный анализ»	17	24			40		81	
Тема 4.1 Кластерный анализ. Основные этапы кластерного анализа.	2	4			4		10	
Тема 4.2 Метрики расстояний. Свойства кластеров. Методы кластерного анализа.	2				4		6	
Тема 4.3 Расстояния между кластерами. Методы группировки кластеров.	2	4			4		10	
Тема 4.4 Иерархические методы кластеризации. Агломеративные методы.	2	4			4		10	

Разделы / Темы	Контактная работа			Часы СР на подготовку кур. р.	Иная СР	Контроль	Всего часов	
	Занятия лекционного типа	Занятия семинарского типа						Контактная работа по кур.р
		Лаб. р	Прак. /сем.					
Тема 4.5 Иерархические дивизимные методы кластеризации.	2	4			4		10	
Тема 4.6 Примеры реализации иерархических методов кластеризации.	2				5		7	
Тема 4.7 Итеративные методы. Метод к-средних.	2	4			5		13	
Тема 4.8 Метод к-средних. Критерии качества кластеризации.	2	4			5		13	
Тема 4.9 Метод нечетких к-средних	1				5		6	
Раздел №5 «Методы классификации»	7	16			40		63	
Тема 5.1 Понятие методов классификации.	1				6		7	
Тема 5.2 Логистическая регрессия.	1	4			6		11	
Тема 5.3 Метрики качества классификации.	1				6		7	
Тема 5.4 Метрики качества классификации. Регуляризация в задаче регрессии.	1	4			6		11	
Тема 5.5 Методы деревьев решений.	1				6		7	
Тема 5.6 Ансамбли деревьев решений: Случайный лес.	1	4			5		10	
Тема 5.7 Ансамбли деревьев решений: Градиентный бустинг.	1	4			5		10	
Индивидуальные консультации при выполнении курсовой работы								
Защита курсовой работы								
Экзамен						36	36	
Итого за 6 семестр	24	40	-			36	180	
Итого по дисциплине	48	80			194	38	360	

ЗАНЯТИЯ СЕМИНАРСКОГО ТИПА для очной формы обучения

Семинарские занятия

Общие рекомендации по подготовке к семинарским занятиям. При подготовке к работе во время проведения занятий семинарского типа следует обратить внимание на следующие моменты: на процесс предварительной подготовки, на работу во время занятия, обработку полученных результатов, исправление полученных замечаний. Предварительная подготовка к учебному занятию семинарского типа заключается в изучении теоретического материала в отведенное для самостоятельной работы время, ознакомление с инструктивными материалами с целью осознания задач занятия. Работа во время проведения занятия семинарского типа включает несколько моментов: а) консультирование обучающихся преподавателями с целью предоставления исчерпывающей информации, необходимой для самостоя-

тельного выполнения предложенных преподавателем задач, б) самостоятельное выполнение заданий согласно обозначенной учебной программой тематики.

5 семестр

Раздел №1. «Введение в статистические методы анализа данных».

Лабораторная работа 1. (4 ч). Графические и аналитические методы выявления аномальных наблюдений.

Порядок выполнения работы:

1. Для предложенной выборки необходимо выполнить проверку аномальных наблюдений в значениях признаков, используя графические и аналитические методы.
2. По результатам проверки требуется сделать обоснованные выводы.

Литература:

Основная

1. Федин, Ф. О. Анализ данных. Часть 1. Подготовка данных к анализу : учебное пособие / Ф. О. Федин, Ф. Ф. Федин. – Москва : Московский городской педагогический университет, 2012. – 204 с. – Текст: электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. — URL: <https://www.iprsmarthop.ru/26444.html>

2. Федин Ф.О. Анализ данных. Часть 2. Инструменты Data Mining: учебное пособие / Федин Ф.О., Федин Ф.Ф. – Москва: Московский городской педагогический университет, 2012. – 308 с. – Текст: электронный // IPR SMART : [сайт]. – URL: <https://www.iprsmarthop.ru/26445.html>

Дополнительная

1. Айвазян С.А. Практикум по многомерным статистическим методам: учебное пособие / Айвазян С.А., Мхитарян В.С., Зехин В.А. – Москва: Евразийский открытый институт, Московский государственный университет экономики, статистики и информатики, 2003. – 76 с. – Текст : электронный // IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/10803.html>

2. Статистический анализ данных, моделирование и исследование вероятностных закономерностей. Компьютерный подход: монография / Б. Ю. Лемешко, С. Б. Лемешко, С. Н. Постовалов, Е. В. Чимитова. – Новосибирск: Новосибирский государственный технический университет, 2011. – 888 с. – ISBN 978-5-7782-1590-0. – Текст: электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/47719.html>

Лабораторная работа 2. (4 ч). Графические и аналитические методы выявления и устранения аномальных наблюдений.

Порядок выполнения работы:

1. Для предложенной выборки необходимо выполнить проверку аномальных наблюдений в значениях признаков, используя изученные критерии.
2. По результатам проверки требуется сделать обоснованные выводы.
3. В случае наличия аномальных наблюдений необходимо предложить и реализовать алгоритмы устранения аномальных наблюдений.

Литература:

Основная

1. Федин, Ф. О. Анализ данных. Часть 1. Подготовка данных к анализу : учебное пособие / Ф. О. Федин, Ф. Ф. Федин. – Москва : Московский городской педагогический университет, 2012. – 204 с. – Текст: электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. — URL: <https://www.iprsmarthop.ru/26444.html>

2. Федин Ф.О. Анализ данных. Часть 2. Инструменты Data Mining: учебное пособие / Федин Ф.О., Федин Ф.Ф. – Москва: Московский городской педагогический университет, 2012. – 308 с. – Текст: электронный // IPR SMART : [сайт]. – URL: <https://www.iprsmarthop.ru/26445.html>

Дополнительная

1. Айвазян С.А. Практикум по многомерным статистическим методам: учебное пособие / Айвазян С.А., Мхитарян В.С., Зехин В.А. – Москва: Евразийский открытый институт, Московский государственный университет экономики, статистики и информатики, 2003. – 76 с. – Текст : электронный // IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/10803.html>

2. Статистический анализ данных, моделирование и исследование вероятностных закономерностей. Компьютерный подход: монография / Б. Ю. Лемешко, С. Б. Лемешко, С. Н. Постовалов, Е. В. Чимитова. – Новосибирск: Новосибирский государственный технический университет, 2011. – 888 с. – ISBN 978-5-7782-1590-0. – Текст: электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/47719.html>

Раздел №2. «Множественный корреляционно-регрессионный анализ. Классическая модель множественной регрессии»

Лабораторная работа 3. Множественный корреляционный анализ (4 ч).

Порядок выполнения работы:

1. Для предложенной выборки необходимо вычислить частные, множественный коэффициенты корреляции, провести пошаговый отбор наиболее информативных переменных.

2. По результатам расчетов требуется сделать обоснованные выводы.

Литература:

Основная

1. Мхитарян В.С. Эконометрика: учебное пособие / Мхитарян В.С., Архипова М.Ю., Сиротин В.П. – Москва : Евразийский открытый институт, 2012. – 224 с. – ISBN 978-5-374-00053-5. – Текст: электронный // IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/11125.html>

2. Федин, Ф. О. Анализ данных. Часть 1. Подготовка данных к анализу : учебное пособие / Ф. О. Федин, Ф. Ф. Федин. – Москва : Московский городской педагогический университет, 2012. – 204 с. – Текст: электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. — URL: <https://www.iprsmarthop.ru/26444.html>

3. Федин Ф.О. Анализ данных. Часть 2. Инструменты Data Mining: учебное пособие / Федин Ф.О., Федин Ф.Ф. – Москва: Московский городской педагогический университет, 2012. – 308 с. – Текст: электронный // IPR SMART : [сайт]. – URL: <https://www.iprsmarthop.ru/26445.html>

Дополнительная литература

1. Статистический анализ данных, моделирование и исследование вероятностных закономерностей. Компьютерный подход: монография / Б. Ю. Лемешко, С. Б. Лемешко, С. Н. Постовалов, Е. В. Чимитова. – Новосибирск: Новосибирский государственный технический университет, 2011. – 888 с. – ISBN 978-5-7782-1590-0. – Текст: электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/47719.html>

2. Тимофеев В.С. Эконометрика: учебник / Тимофеев В.С., Фаддеенков А.В., Щеколдин В.Ю.. – Новосибирск: Новосибирский государственный технический университет, 2015. – 352 с. – ISBN 978-5-7782-2658-6. – Текст: электронный // IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/91744.html>

3. Цильковский И.А. Методы анализа знаний и данных: конспект лекций / Цильковский И.А., Волкова В.М.. – Новосибирск: Новосибирский государственный технический университет, 2010. – 68 с. – ISBN 978-57782-1377-7. – Текст: электронный // IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/45385.html>

Лабораторная работа 4. Построение и проверка адекватности классической модели множественной регрессии (4 ч).

Порядок выполнения работы:

1. Из предложенного статистического сборника требуется выбрать переменные для моделирования, обосновать свой выбор.

2. На основе сформированной выборки наблюдений необходимо построить эмпирическое уравнение множественной линейной регрессии для выбранного набора зависимой и объясняющих переменных.

3. По результатам расчетов требуется провести проверку адекватности построенного уравнения регрессии, построить доверительные интервалы для теоретических коэффициентов.

Литература:

Основная

1. Мхитарян В.С. Эконометрика: учебное пособие / Мхитарян В.С., Архипова М.Ю., Сиротин В.П. – Москва : Евразийский открытый институт, 2012. – 224 с. – ISBN 978-5-374-00053-5. – Текст: электронный // IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/11125.html>

2. Федин, Ф. О. Анализ данных. Часть 1. Подготовка данных к анализу : учебное пособие / Ф. О. Федин, Ф. Ф. Федин. – Москва : Московский городской педагогический университет, 2012. – 204 с. – Текст: электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. — URL: <https://www.iprsmarthop.ru/26444.html>

3. Федин Ф.О. Анализ данных. Часть 2. Инструменты Data Mining: учебное пособие / Федин Ф.О., Федин Ф.Ф. – Москва: Московский городской педагогический университет, 2012. – 308 с. – Текст: электронный // IPR SMART : [сайт]. – URL: <https://www.iprsmarthop.ru/26445.html>

Дополнительная литература

1. Статистический анализ данных, моделирование и исследование вероятностных закономерностей. Компьютерный подход: монография / Б. Ю. Лемешко, С. Б. Лемешко, С. Н. Постовалов, Е. В.

Чимитова. – Новосибирск: Новосибирский государственный технический университет, 2011. – 888 с. – ISBN 978-5-7782-1590-0. – Текст: электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/47719.html>

2. Тимофеев В.С. Эконометрика: учебник / Тимофеев В.С., Фаддеенков А.В., Щеколдин В.Ю.. – Новосибирск: Новосибирский государственный технический университет, 2015. – 352 с. – ISBN 978-5-7782-2658-6. – Текст: электронный // IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/91744.html>

3. Цильковский И.А. Методы анализа знаний и данных: конспект лекций / Цильковский И.А., Волкова В.М.. – Новосибирск: Новосибирский государственный технический университет, 2010. – 68 с. – ISBN 978-57782-1377-7. – Текст: электронный // IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/45385.html>

Лабораторная работа 5. Построение и проверка адекватности классической модели множественной регрессии (4 ч).

Порядок выполнения работы:

1. Из предложенного статистического сборника требуется выбрать переменные для моделирования, обосновать свой выбор.

2. На основе сформированной выборки наблюдений необходимо построить эмпирическое уравнение множественной линейной регрессии для выбранного набора зависимой и объясняющих переменных, провести проверку адекватности построенного уравнения регрессии.

3. По результатам расчетов требуется сделать обоснованные выводы.

Литература:

Основная

1. Мхитарян В.С. Эконометрика: учебное пособие / Мхитарян В.С., Архипова М.Ю., Сиротин В.П.. – Москва : Евразийский открытый институт, 2012. – 224 с. – ISBN 978-5-374-00053-5. – Текст: электронный // IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/11125.html>

2. Федин, Ф. О. Анализ данных. Часть 1. Подготовка данных к анализу : учебное пособие / Ф. О. Федин, Ф. Ф. Федин. – Москва : Московский городской педагогический университет, 2012. – 204 с. – Текст: электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. — URL: <https://www.iprsmarthop.ru/26444.html>

3. Федин Ф.О. Анализ данных. Часть 2. Инструменты Data Mining: учебное пособие / Федин Ф.О., Федин Ф.Ф. – Москва: Московский городской педагогический университет, 2012. – 308 с. – Текст: электронный // IPR SMART : [сайт]. – URL: <https://www.iprsmarthop.ru/26445.html>

Дополнительная литература

1. Статистический анализ данных, моделирование и исследование вероятностных закономерностей. Компьютерный подход: монография / Б. Ю. Лемешко, С. Б. Лемешко, С. Н. Постовалов, Е. В. Чимитова. – Новосибирск: Новосибирский государственный технический университет, 2011. – 888 с. – ISBN 978-5-7782-1590-0. – Текст: электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/47719.html>

2. Тимофеев В.С. Эконометрика: учебник / Тимофеев В.С., Фаддеенков А.В., Щеколдин В.Ю.. – Новосибирск: Новосибирский государственный технический университет, 2015. – 352 с. – ISBN 978-5-7782-2658-6. – Текст: электронный // IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/91744.html>

3. Цильковский И.А. Методы анализа знаний и данных: конспект лекций / Цильковский И.А., Волкова В.М.. – Новосибирск: Новосибирский государственный технический университет, 2010. – 68 с. – ISBN 978-57782-1377-7. – Текст: электронный // IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/45385.html>

Лабораторная работа 6. Построение регрессионной модели с фиктивными переменными (4 ч).

Порядок выполнения работы:

1. На основе сформированной выборки наблюдений необходимо:
– построить эмпирическое уравнение множественной линейной регрессии для выбранного набора зависимой и объясняющих переменных, предварительно выполнив процедуру кодирования категориальных переменных;
– провести проверку адекватности построенного уравнения регрессии.

2. По результатам расчетов требуется сделать обоснованные выводы.

Литература:

Основная

4. Мхитарян В.С. Эконометрика: учебное пособие / Мхитарян В.С., Архипова М.Ю., Сиротин В.П. – Москва : Евразийский открытый институт, 2012. – 224 с. – ISBN 978-5-374-00053-5. – Текст: электронный // IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/11125.html>

5. Федин, Ф. О. Анализ данных. Часть 1. Подготовка данных к анализу : учебное пособие / Ф. О. Федин, Ф. Ф. Федин. – Москва : Московский городской педагогический университет, 2012. – 204 с. – Текст: электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. — URL: <https://www.iprsmarthop.ru/26444.html>

6. Федин Ф.О. Анализ данных. Часть 2. Инструменты Data Mining: учебное пособие / Федин Ф.О., Федин Ф.Ф. – Москва: Московский городской педагогический университет, 2012. – 308 с. – Текст: электронный // IPR SMART : [сайт]. – URL: <https://www.iprsmarthop.ru/26445.html>

Дополнительная литература

4. Статистический анализ данных, моделирование и исследование вероятностных закономерностей. Компьютерный подход: монография / Б. Ю. Лемешко, С. Б. Лемешко, С. Н. Постовалов, Е. В. Чимитова. – Новосибирск: Новосибирский государственный технический университет, 2011. – 888 с. – ISBN 978-5-7782-1590-0. – Текст: электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/47719.html>

5. Тимофеев В.С. Эконометрика: учебник / Тимофеев В.С., Фаддеенков А.В., Щеколдин В.Ю.. – Новосибирск: Новосибирский государственный технический университет, 2015. – 352 с. – ISBN 978-5-7782-2658-6. – Текст: электронный // IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/91744.html>

6. Цильковский И.А. Методы анализа знаний и данных: конспект лекций / Цильковский И.А., Волкова В.М.. – Новосибирск: Новосибирский государственный технический университет, 2010. – 68 с. – ISBN 978-57782-1377-7. – Текст: электронный // IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/45385.html>

Лабораторная работа 7. Проверка мультиколлинеарности. Устранение мультиколлинеарности (4 ч).

Порядок выполнения работы:

1. Из предложенного статистического сборника требуется выбрать переменные для моделирования, обосновать свой выбор.

2. На основе сформированной выборки наблюдений необходимо проверить наличие мультиколлинеарности переменных, в случае наличия мультиколлинеарности предложить и реализовать алгоритм устранения мультиколлинеарности.

3. По результатам расчетов требуется сделать обоснованные выводы.

Литература:

Основная

1. Мхитарян В.С. Эконометрика: учебное пособие / Мхитарян В.С., Архипова М.Ю., Сиротин В.П. – Москва : Евразийский открытый институт, 2012. – 224 с. – ISBN 978-5-374-00053-5. – Текст: электронный // IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/11125.html>

2. Федин, Ф. О. Анализ данных. Часть 1. Подготовка данных к анализу : учебное пособие / Ф. О. Федин, Ф. Ф. Федин. – Москва : Московский городской педагогический университет, 2012. – 204 с. – Текст: электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. — URL: <https://www.iprsmarthop.ru/26444.html>

3. Федин Ф.О. Анализ данных. Часть 2. Инструменты Data Mining: учебное пособие / Федин Ф.О., Федин Ф.Ф. – Москва: Московский городской педагогический университет, 2012. – 308 с. – Текст: электронный // IPR SMART : [сайт]. – URL: <https://www.iprsmarthop.ru/26445.html>

Дополнительная литература

1. Статистический анализ данных, моделирование и исследование вероятностных закономерностей. Компьютерный подход: монография / Б. Ю. Лемешко, С. Б. Лемешко, С. Н. Постовалов, Е. В. Чимитова. – Новосибирск: Новосибирский государственный технический университет, 2011. – 888 с. – ISBN 978-5-7782-1590-0. – Текст: электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/47719.html>

2. Тимофеев В.С. Эконометрика: учебник / Тимофеев В.С., Фаддеенков А.В., Щеколдин В.Ю.. – Новосибирск: Новосибирский государственный технический университет, 2015. – 352 с. – ISBN 978-5-7782-2658-6. – Текст: электронный // IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/91744.html>

3. Цильковский И.А. Методы анализа знаний и данных: конспект лекций / Цильковский И.А., Волкова В.М.. – Новосибирск: Новосибирский государственный технический университет, 2010. – 68 с. – ISBN 978-57782-1377-7. – Текст: электронный // IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/45385.html>

Раздел №3. «Обобщенная линейная модель множественной регрессии. Регрессионные модели с автокоррелированными остатками»

Лабораторная работа 8. Критерии проверки гетероскедастичности в остатках (4 ч).

Порядок выполнения работы:

1. На основе сформированной выборки наблюдений необходимо построить эмпирическое уравнение множественной линейной регрессии для выбранного набора зависимой и объясняющих переменных, провести проверку адекватности построенного уравнения регрессии.

2. Для построенного эмпирического уравнения регрессии требуется провести проверку остатков на наличие гетероскедастичности, используя графический анализ, аналитические критерии проверки гетероскедастичности.

3. По результатам расчетов требуется сделать обоснованные выводы.

Литература:

Основная

1. Мхитарян В.С. Эконометрика: учебное пособие / Мхитарян В.С., Архипова М.Ю., Сиротин В.П. – Москва : Евразийский открытый институт, 2012. – 224 с. – ISBN 978-5-374-00053-5. – Текст: электронный // IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/11125.html>

2. Федин, Ф. О. Анализ данных. Часть 1. Подготовка данных к анализу : учебное пособие / Ф. О. Федин, Ф. Ф. Федин. – Москва : Московский городской педагогический университет, 2012. – 204 с. – Текст: электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. — URL: <https://www.iprsmarthop.ru/26444.html>

3. Федин Ф.О. Анализ данных. Часть 2. Инструменты Data Mining: учебное пособие / Федин Ф.О., Федин Ф.Ф. – Москва: Московский городской педагогический университет, 2012. – 308 с. – Текст: электронный // IPR SMART : [сайт]. – URL: <https://www.iprsmarthop.ru/26445.html>

Дополнительная литература

1. Статистический анализ данных, моделирование и исследование вероятностных закономерностей. Компьютерный подход: монография / Б. Ю. Лемешко, С. Б. Лемешко, С. Н. Постовалов, Е. В. Чимитова. – Новосибирск: Новосибирский государственный технический университет, 2011. – 888 с. – ISBN 978-5-7782-1590-0. – Текст: электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/47719.html>

2. Тимофеев В.С. Эконометрика: учебник / Тимофеев В.С., Фаддеенков А.В., Щеколдин В.Ю.. – Новосибирск: Новосибирский государственный технический университет, 2015. – 352 с. – ISBN 978-5-7782-2658-6. – Текст: электронный // IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/91744.html>

3. Цильковский И.А. Методы анализа знаний и данных: конспект лекций / Цильковский И.А., Волкова В.М.. – Новосибирск: Новосибирский государственный технический университет, 2010. – 68 с. – ISBN 978-57782-1377-7. – Текст: электронный // IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/45385.html>

Лабораторная работа 9. Взвешенный МНК для случая гетероскедастичности в остатках (4 ч).

Порядок выполнения работы:

1. На основе сформированной выборки наблюдений необходимо построить эмпирическое уравнение множественной линейной регрессии для выбранного набора зависимой и объясняющих переменных, провести проверку адекватности построенного уравнения регрессии.

2. Для построенного эмпирического уравнения регрессии требуется провести проверку остатков на наличие гетероскедастичности, используя графический анализ, аналитические критерии проверки гетероскедастичности.

3. В случае наличия гетероскедастичности предложить и реализовать алгоритм ее устранения, проверить остатки модели на наличие гетероскедастичности.

4. По результатам расчетов требуется сделать обоснованные выводы.

Литература:

Основная

1. Мхитарян В.С. Эконометрика: учебное пособие / Мхитарян В.С., Архипова М.Ю., Сиротин В.П. – Москва : Евразийский открытый институт, 2012. – 224 с. – ISBN 978-5-374-00053-5. – Текст: электронный // IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/11125.html>

2. Федин, Ф. О. Анализ данных. Часть 1. Подготовка данных к анализу : учебное пособие / Ф. О. Федин, Ф. Ф. Федин. – Москва : Московский городской педагогический университет, 2012. – 204 с.

– Текст: электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. — URL: <https://www.iprsmarthop.ru/26444.html>

3. Федин Ф.О. Анализ данных. Часть 2. Инструменты Data Mining: учебное пособие / Федин Ф.О., Федин Ф.Ф. – Москва: Московский городской педагогический университет, 2012. – 308 с. – Текст: электронный // IPR SMART : [сайт]. – URL: <https://www.iprsmarthop.ru/26445.html>

Дополнительная литература

1. Статистический анализ данных, моделирование и исследование вероятностных закономерностей. Компьютерный подход: монография / Б. Ю. Лемешко, С. Б. Лемешко, С. Н. Постовалов, Е. В. Чимитова. – Новосибирск: Новосибирский государственный технический университет, 2011. – 888 с. – ISBN 978-5-7782-1590-0. – Текст: электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/47719.html>

2. Тимофеев В.С. Эконометрика: учебник / Тимофеев В.С., Фаддеев А.В., Щеколдин В.Ю.. – Новосибирск: Новосибирский государственный технический университет, 2015. – 352 с. – ISBN 978-5-7782-2658-6. – Текст: электронный // IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/91744.html>

3. Цильковский И.А. Методы анализа знаний и данных: конспект лекций / Цильковский И.А., Волкова В.М.. – Новосибирск: Новосибирский государственный технический университет, 2010. – 68 с. – ISBN 978-57782-1377-7. – Текст: электронный // IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/45385.html>

Лабораторная работа 10. Критерии проверки и устранения автокорреляции в остатках (4 ч).

Порядок выполнения работы:

1. На основе сформированной выборки наблюдений необходимо построить эмпирическое уравнение множественной линейной регрессии для выбранного набора зависимой и объясняющих переменных, провести проверку адекватности построенного уравнения регрессии.

2. Для построенного эмпирического уравнения регрессии требуется провести проверку остатков на наличие автокорреляции, используя графический анализ, аналитические критерии проверки автокорреляции.

3. В случае наличия автокорреляции предложить и реализовать алгоритм ее устранения, проверить остатки модели на наличие автокорреляции.

4. По результатам расчетов требуется сделать обоснованные выводы.

Литература:

Основная

1. Мхитарян В.С. Эконометрика: учебное пособие / Мхитарян В.С., Архипова М.Ю., Сиротин В.П.. – Москва: Евразийский открытый институт, 2012. – 224 с. – ISBN 978-5-374-00053-5. – Текст: электронный // IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/11125.html>

2. Федин, Ф. О. Анализ данных. Часть 1. Подготовка данных к анализу : учебное пособие / Ф. О. Федин, Ф. Ф. Федин. – Москва : Московский городской педагогический университет, 2012. – 204 с. – Текст: электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. — URL: <https://www.iprsmarthop.ru/26444.html>

3. Федин Ф.О. Анализ данных. Часть 2. Инструменты Data Mining: учебное пособие / Федин Ф.О., Федин Ф.Ф. – Москва: Московский городской педагогический университет, 2012. – 308 с. – Текст: электронный // IPR SMART : [сайт]. – URL: <https://www.iprsmarthop.ru/26445.html>

Дополнительная литература

1. Статистический анализ данных, моделирование и исследование вероятностных закономерностей. Компьютерный подход: монография / Б. Ю. Лемешко, С. Б. Лемешко, С. Н. Постовалов, Е. В. Чимитова. – Новосибирск: Новосибирский государственный технический университет, 2011. – 888 с. – ISBN 978-5-7782-1590-0. – Текст: электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/47719.html>

2. Тимофеев В.С. Эконометрика: учебник / Тимофеев В.С., Фаддеев А.В., Щеколдин В.Ю.. – Новосибирск: Новосибирский государственный технический университет, 2015. – 352 с. – ISBN 978-5-7782-2658-6. – Текст: электронный // IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/91744.html>

3. Цильковский И.А. Методы анализа знаний и данных: конспект лекций / Цильковский И.А., Волкова В.М.. – Новосибирск: Новосибирский государственный технический университет, 2010. – 68 с. – ISBN 978-57782-1377-7. – Текст: электронный // IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/45385.html>

Раздел №4. «Кластерный анализ».

Лабораторная работа 11. Иерархические агломеративные методы кластеризации (4 ч).

Порядок выполнения работы:

1. По предложенной выборке наблюдений требуется провести кластеризацию, используя иерархический агломеративный алгоритм кластеризации. В качестве расстояния между элементами использовать евклидово расстояние, в качестве расстояния между кластерами – расстояние, измеряемое по принципу ближайшего соседа (одиночной связи).

2. По результатам расчетов требуется провести анализ полученного разбиения.

Литература:

Основная

1. Федин, Ф. О. Анализ данных. Часть 1. Подготовка данных к анализу : учебное пособие / Ф. О. Федин, Ф. Ф. Федин. – Москва : Московский городской педагогический университет, 2012. – 204 с. – Текст: электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. — URL: <https://www.iprsmarthop.ru/26444.html>

2. Федин Ф.О. Анализ данных. Часть 2. Инструменты Data Mining: учебное пособие / Федин Ф.О., Федин Ф.Ф. – Москва: Московский городской педагогический университет, 2012. – 308 с. – Текст: электронный // IPR SMART : [сайт]. – URL: <https://www.iprsmarthop.ru/26445.html>

Дополнительная литература

1. Статистический анализ данных, моделирование и исследование вероятностных закономерностей. Компьютерный подход: монография / Б. Ю. Лемешко, С. Б. Лемешко, С. Н. Постовалов, Е. В. Чимитова. – Новосибирск: Новосибирский государственный технический университет, 2011. – 888 с. – ISBN 978-5-7782-1590-0. – Текст: электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/47719.html>

2. Цильковский И.А. Методы анализа знаний и данных: конспект лекций / Цильковский И.А., Волкова В.М.. – Новосибирск: Новосибирский государственный технический университет, 2010. – 68 с. – ISBN 978-57782-1377-7. – Текст: электронный // IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/45385.html>

Лабораторная работа 12. Иерархические агломеративные методы кластеризации (4 ч).

Порядок выполнения работы:

1. По предложенной выборке наблюдений требуется провести кластеризацию, используя иерархический агломеративный алгоритм кластеризации. В качестве расстояния между элементами использовать евклидово расстояние, в качестве расстояния между кластерами – расстояние, измеряемое по принципу средней связи.

2. По результатам расчетов требуется провести анализ полученного разбиения.

Литература:

Основная

1. Федин, Ф. О. Анализ данных. Часть 1. Подготовка данных к анализу : учебное пособие / Ф. О. Федин, Ф. Ф. Федин. – Москва : Московский городской педагогический университет, 2012. – 204 с. – Текст: электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. — URL: <https://www.iprsmarthop.ru/26444.html>

2. Федин Ф.О. Анализ данных. Часть 2. Инструменты Data Mining: учебное пособие / Федин Ф.О., Федин Ф.Ф. – Москва: Московский городской педагогический университет, 2012. – 308 с. – Текст: электронный // IPR SMART : [сайт]. – URL: <https://www.iprsmarthop.ru/26445.html>

Дополнительная литература

1. Статистический анализ данных, моделирование и исследование вероятностных закономерностей. Компьютерный подход: монография / Б. Ю. Лемешко, С. Б. Лемешко, С. Н. Постовалов, Е. В. Чимитова. – Новосибирск: Новосибирский государственный технический университет, 2011. – 888 с. – ISBN 978-5-7782-1590-0. – Текст: электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/47719.html>

2. Цильковский И.А. Методы анализа знаний и данных: конспект лекций / Цильковский И.А., Волкова В.М.. – Новосибирск: Новосибирский государственный технический университет, 2010. – 68 с. – ISBN 978-57782-1377-7. – Текст: электронный // IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/45385.html>

Лабораторная работа 13. Иерархические дивизимные методы кластеризации (4 ч).

Порядок выполнения работы:

1. По предложенной выборке наблюдений требуется провести кластеризацию, используя иерархический дивизимный алгоритм кластеризации. В качестве расстояния между элементами использовать евклидово расстояние, в качестве расстояния между кластерами – расстояние, измеряемое по принципу средней связи.

2. По результатам расчетов требуется провести анализ полученного разбиения.

Литература:

Основная

1. Федин, Ф. О. Анализ данных. Часть 1. Подготовка данных к анализу : учебное пособие / Ф. О. Федин, Ф. Ф. Федин. – Москва : Московский городской педагогический университет, 2012. – 204 с. – Текст: электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. — URL: <https://www.iprsmarthop.ru/26444.html>

2. Федин Ф.О. Анализ данных. Часть 2. Инструменты Data Mining: учебное пособие / Федин Ф.О., Федин Ф.Ф. – Москва: Московский городской педагогический университет, 2012. – 308 с. – Текст: электронный // IPR SMART : [сайт]. – URL: <https://www.iprsmarthop.ru/26445.html>

Дополнительная литература

1. Статистический анализ данных, моделирование и исследование вероятностных закономерностей. Компьютерный подход: монография / Б. Ю. Лемешко, С. Б. Лемешко, С. Н. Постовалов, Е. В. Чимитова. – Новосибирск: Новосибирский государственный технический университет, 2011. – 888 с. – ISBN 978-5-7782-1590-0. – Текст: электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/47719.html>

2. Цильковский И.А. Методы анализа знаний и данных: конспект лекций / Цильковский И.А., Волкова В.М.. – Новосибирск: Новосибирский государственный технический университет, 2010. – 68 с. – ISBN 978-57782-1377-7. – Текст: электронный // IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/45385.html>

Лабораторная работа 14. Метод k-средних (4 ч).

Порядок выполнения работы:

1. По предложенной выборке требуется провести кластеризацию, используя метод k-средних для $k=3$. В качестве расстояния между элементами использовать евклидово расстояние. В качестве начальных центроидов выбрать три наиболее отдаленных друг от друга объекта.

2. На каждом шаге разбиения требуется вычислить основные функционалы качества.

3. По результатам расчетов требуется провести анализ полученного разбиения.

Литература:

Основная

1. Федин, Ф. О. Анализ данных. Часть 1. Подготовка данных к анализу : учебное пособие / Ф. О. Федин, Ф. Ф. Федин. – Москва : Московский городской педагогический университет, 2012. – 204 с. – Текст: электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. — URL: <https://www.iprsmarthop.ru/26444.html>

2. Федин Ф.О. Анализ данных. Часть 2. Инструменты Data Mining: учебное пособие / Федин Ф.О., Федин Ф.Ф. – Москва: Московский городской педагогический университет, 2012. – 308 с. – Текст: электронный // IPR SMART : [сайт]. – URL: <https://www.iprsmarthop.ru/26445.html>

Дополнительная литература

1. Статистический анализ данных, моделирование и исследование вероятностных закономерностей. Компьютерный подход: монография / Б. Ю. Лемешко, С. Б. Лемешко, С. Н. Постовалов, Е. В. Чимитова. – Новосибирск: Новосибирский государственный технический университет, 2011. – 888 с. – ISBN 978-5-7782-1590-0. – Текст: электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/47719.html>

2. Цильковский И.А. Методы анализа знаний и данных: конспект лекций / Цильковский И.А., Волкова В.М.. – Новосибирск: Новосибирский государственный технический университет, 2010. – 68 с. – ISBN 978-57782-1377-7. – Текст: электронный // IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/45385.html>

Лабораторная работа 15. Выбор оптимальной кластеризации (4 ч).

Порядок выполнения работы:

1. По предложенной выборке показателей социально-экономического развития субъектов РФ требуется построить оптимальное разбиение субъектов, используя метод k-средних:

– провести унификацию исходных данных;

– провести кластеризацию, используя метод k-средних, для $k=3$;

– провести кластеризацию, используя метод k-средних, для $k=4$;

– провести кластеризацию, используя метод k -средних, для $k=5$.

2. Используя метод Локтя, выбрать оптимальное разбиение.

Литература:

Основная

1. Федин, Ф. О. Анализ данных. Часть 1. Подготовка данных к анализу : учебное пособие / Ф. О. Федин, Ф. Ф. Федин. – Москва : Московский городской педагогический университет, 2012. – 204 с. – Текст: электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. — URL: <https://www.iprsmarthop.ru/26444.html>

2. Федин Ф.О. Анализ данных. Часть 2. Инструменты Data Mining: учебное пособие / Федин Ф.О., Федин Ф.Ф. – Москва: Московский городской педагогический университет, 2012. – 308 с. – Текст: электронный // IPR SMART : [сайт]. – URL: <https://www.iprsmarthop.ru/26445.html>

Дополнительная литература

1. Статистический анализ данных, моделирование и исследование вероятностных закономерностей. Компьютерный подход: монография / Б. Ю. Лемешко, С. Б. Лемешко, С. Н. Постовалов, Е. В. Чимитова. – Новосибирск: Новосибирский государственный технический университет, 2011. – 888 с. – ISBN 978-5-7782-1590-0. – Текст: электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/47719.html>

2. Цильковский И.А. Методы анализа знаний и данных: конспект лекций / Цильковский И.А., Волкова В.М.. – Новосибирск: Новосибирский государственный технический университет, 2010. – 68 с. – ISBN 978-57782-1377-7. – Текст: электронный // IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/45385.html>

Лабораторная работа 16. Метод нечетких k -средних (4 ч).

Порядок выполнения работы:

1. По предложенной выборке требуется провести кластеризацию, используя метод нечетких k -средних для $k=3$. В качестве расстояния между элементами использовать евклидово расстояние. В качестве начальных центроидов выбрать три наиболее отдаленных друг от друга объекта.

2. На каждом шаге разбиения требуется вычислить основные функционалы качества.

3. По результатам расчетов требуется провести анализ полученного разбиения.

Литература:

Основная

1. Федин, Ф. О. Анализ данных. Часть 1. Подготовка данных к анализу : учебное пособие / Ф. О. Федин, Ф. Ф. Федин. – Москва : Московский городской педагогический университет, 2012. – 204 с. – Текст: электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. — URL: <https://www.iprsmarthop.ru/26444.html>

2. Федин Ф.О. Анализ данных. Часть 2. Инструменты Data Mining: учебное пособие / Федин Ф.О., Федин Ф.Ф. – Москва: Московский городской педагогический университет, 2012. – 308 с. – Текст: электронный // IPR SMART : [сайт]. – URL: <https://www.iprsmarthop.ru/26445.html>

Дополнительная литература

1. Статистический анализ данных, моделирование и исследование вероятностных закономерностей. Компьютерный подход: монография / Б. Ю. Лемешко, С. Б. Лемешко, С. Н. Постовалов, Е. В. Чимитова. – Новосибирск: Новосибирский государственный технический университет, 2011. – 888 с. – ISBN 978-5-7782-1590-0. – Текст: электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/47719.html>

2. Цильковский И.А. Методы анализа знаний и данных: конспект лекций / Цильковский И.А., Волкова В.М.. – Новосибирск: Новосибирский государственный технический университет, 2010. – 68 с. – ISBN 978-57782-1377-7. – Текст: электронный // IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/45385.html>

Раздел №5. «Методы классификации»

Лабораторная работа 17. Построение логистической регрессии (4 ч).

Порядок выполнения работы:

1. На основе сформированной выборки наблюдений необходимо построить уравнение логистической регрессии для выбранного набора зависимой и объясняющих переменных.

2. По результатам расчетов провести проверку качества построенной модели.

Литература:

Основная

1. Федин, Ф. О. Анализ данных. Часть 1. Подготовка данных к анализу : учебное пособие / Ф. О. Федин, Ф. Ф. Федин. – Москва : Московский городской педагогический университет, 2012. – 204 с.

– Текст: электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. — URL: <https://www.iprsmarthop.ru/26444.html>

2. Федин Ф.О. Анализ данных. Часть 2. Инструменты Data Mining: учебное пособие / Федин Ф.О., Федин Ф.Ф. – Москва: Московский городской педагогический университет, 2012. – 308 с. – Текст: электронный // IPR SMART : [сайт]. – URL: <https://www.iprsmarthop.ru/26445.html>

Дополнительная литература

1. Статистический анализ данных, моделирование и исследование вероятностных закономерностей. Компьютерный подход: монография / Б. Ю. Лемешко, С. Б. Лемешко, С. Н. Постовалов, Е. В. Чимитова. – Новосибирск: Новосибирский государственный технический университет, 2011. – 888 с. – ISBN 978-5-7782-1590-0. – Текст: электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/47719.html>

2. Цильковский И.А. Методы анализа знаний и данных: конспект лекций / Цильковский И.А., Волкова В.М.. – Новосибирск: Новосибирский государственный технический университет, 2010. – 68 с. – ISBN 978-57782-1377-7. – Текст: электронный // IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/45385.html>

Лабораторная работа 18. Построение деревьев решений (4 ч).

Порядок выполнения работы:

1. На основе сформированной выборки наблюдений необходимо построить модель дерева решений для выбранного набора зависимой и объясняющих переменных.
2. По результатам расчетов провести проверку качества построенной модели.

Литература:

Основная

1. Федин, Ф. О. Анализ данных. Часть 1. Подготовка данных к анализу : учебное пособие / Ф. О. Федин, Ф. Ф. Федин. – Москва : Московский городской педагогический университет, 2012. – 204 с. – Текст: электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. — URL: <https://www.iprsmarthop.ru/26444.html>

2. Федин Ф.О. Анализ данных. Часть 2. Инструменты Data Mining: учебное пособие / Федин Ф.О., Федин Ф.Ф. – Москва: Московский городской педагогический университет, 2012. – 308 с. – Текст: электронный // IPR SMART : [сайт]. – URL: <https://www.iprsmarthop.ru/26445.html>

Дополнительная литература

1. Статистический анализ данных, моделирование и исследование вероятностных закономерностей. Компьютерный подход: монография / Б. Ю. Лемешко, С. Б. Лемешко, С. Н. Постовалов, Е. В. Чимитова. – Новосибирск: Новосибирский государственный технический университет, 2011. – 888 с. – ISBN 978-5-7782-1590-0. – Текст: электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/47719.html>

2. Цильковский И.А. Методы анализа знаний и данных: конспект лекций / Цильковский И.А., Волкова В.М.. – Новосибирск: Новосибирский государственный технический университет, 2010. – 68 с. – ISBN 978-57782-1377-7. – Текст: электронный // IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/45385.html>

Лабораторная работа 19. Построение деревьев решений (4 ч).

Порядок выполнения работы:

1. На основе сформированной выборки наблюдений необходимо построить модель дерева решений для выбранного набора зависимой и объясняющих переменных при различных параметрах моделирования.
2. По результатам расчетов провести проверку качества построенной модели.

Литература:

Основная

1. Федин, Ф. О. Анализ данных. Часть 1. Подготовка данных к анализу : учебное пособие / Ф. О. Федин, Ф. Ф. Федин. – Москва : Московский городской педагогический университет, 2012. – 204 с. – Текст: электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. — URL: <https://www.iprsmarthop.ru/26444.html>

2. Федин Ф.О. Анализ данных. Часть 2. Инструменты Data Mining: учебное пособие / Федин Ф.О., Федин Ф.Ф. – Москва: Московский городской педагогический университет, 2012. – 308 с. – Текст: электронный // IPR SMART : [сайт]. – URL: <https://www.iprsmarthop.ru/26445.html>

Дополнительная литература

1. Статистический анализ данных, моделирование и исследование вероятностных закономерностей. Компьютерный подход: монография / Б. Ю. Лемешко, С. Б. Лемешко, С. Н. Постовалов, Е. В.

Чимитова. – Новосибирск: Новосибирский государственный технический университет, 2011. – 888 с. – ISBN 978-5-7782-1590-0. – Текст: электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/47719.html>

2. Цильковский И.А. Методы анализа знаний и данных: конспект лекций / Цильковский И.А., Волкова В.М.. – Новосибирск: Новосибирский государственный технический университет, 2010. – 68 с. – ISBN 978-57782-1377-7. – Текст: электронный // IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/45385.html>

Лабораторная работа 20. Оценка качества классификационных моделей (4 ч).

Порядок выполнения работы:

1. Для предложенной выборки наблюдений необходимо построить модели классификации при различных параметрах моделирования.
2. На основе проведенных расчетов провести проверку качества построенных моделей, выбрать оптимальную модель классификации, обосновать свой выбор.

Литература:

Основная

1. Федин, Ф. О. Анализ данных. Часть 1. Подготовка данных к анализу : учебное пособие / Ф. О. Федин, Ф. Ф. Федин. – Москва : Московский городской педагогический университет, 2012. – 204 с. – Текст: электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. — URL: <https://www.iprsmarthop.ru/26444.html>

2. Федин Ф.О. Анализ данных. Часть 2. Инструменты Data Mining: учебное пособие / Федин Ф.О., Федин Ф.Ф. – Москва: Московский городской педагогический университет, 2012. – 308 с. – Текст: электронный // IPR SMART : [сайт]. – URL: <https://www.iprsmarthop.ru/26445.html>

Дополнительная литература

1. Статистический анализ данных, моделирование и исследование вероятностных закономерностей. Компьютерный подход: монография / Б. Ю. Лемешко, С. Б. Лемешко, С. Н. Постовалов, Е. В. Чимитова. – Новосибирск: Новосибирский государственный технический университет, 2011. – 888 с. – ISBN 978-5-7782-1590-0. – Текст: электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/47719.html>

2. Цильковский И.А. Методы анализа знаний и данных: конспект лекций / Цильковский И.А., Волкова В.М.. – Новосибирск: Новосибирский государственный технический университет, 2010. – 68 с. – ISBN 978-57782-1377-7. – Текст: электронный // IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/45385.html>

РАЗДЕЛ 5. ОБРАЗОВАТЕЛЬНЫЕ ТЕХНОЛОГИИ

**Интерактивные образовательные технологии,
используемые на аудиторных лабораторных занятиях**

Очная форма обучения

Наименование разделов, тем	Используемые образовательные технологии	Часы
<p>Раздел №1 «Введение в статистические методы анализа данных» Тема 1.1 Введение в статистические методы анализа данных. Тема 1.2 Методы первичной обработки и представления данных. Тема 1.3 Методы выявления и устранения аномальных наблюдений.</p>	Обсуждение решений профессионально-ориентированных заданий и задач; обсуждение и анализ решения кейсов	8
<p>Раздел №2 «Множественный корреляционно-регрессионный анализ. Классическая модель множественной регрессии» Тема 2.1 Корреляционно-регрессионный анализ. Модель множественной регрессии. Тема 2.2 Фиктивные переменные. Критерий Чоу проверки</p>	Обсуждение решений профессионально-ориентированных заданий и задач; обсуждение и анализ решения кейсов	20

<p>однородности. Тема 2.3 Множественные коэффициенты корреляции. Тема 2.4 Статистический анализ адекватности модели множественной регрессии. Тема 2.5 Мультиколлинеарность факторов.</p>		
<p>Раздел №3 «Обобщенная линейная модель множественной регрессии. Регрессионные модели с автокоррелированными остатками» Тема 3.1 Обобщенная линейная модель множественной регрессии. Обобщенный МНК Тема 3.2 Гетероскедастичность. Методы выявления гетероскедастичности. Тема 3.3 Взвешенный МНК. Устранение гетероскедастичности. Тема 3.4 Регрессионные модели с автокоррелированными остатками. Тема 3.5 Устранение автокорреляции</p>	<p>Обсуждение решений профессионально-ориентированных заданий и задач; обсуждение и анализ решения кейсов</p>	12
<p>Раздел №4 «Кластерный анализ» Тема 4.1 Кластерный анализ. Основные этапы кластерного анализа. Тема 4.2 Метрики расстояний. Свойства кластеров. Методы кластерного анализа. Тема 4.3 Расстояния между кластерами. Методы группировки кластеров. Тема 4.4 Иерархические методы кластеризации. Агломеративные методы. Тема 4.5 Иерархические дивизимные методы кластеризации. Тема 4.6 Примеры реализации иерархических методов кластеризации. Тема 4.7 Итеративные методы. Метод к-средних. Тема 4.8 Метод к-средних. Критерии качества кластеризации. Тема 4.9 Метод нечетких к-средних</p>	<p>Обсуждение решений профессионально-ориентированных заданий и задач; обсуждение и анализ решения кейсов</p>	24
<p>Раздел №5 «Методы классификации» Тема 5.1 Понятие методов классификации. Тема 5.2 Логистическая регрессия. Тема 5.3 Метрики качества классификации. Тема 5.4 Метрики качества классификации. Регуляризация в задаче регрессии. Тема 5.5 Методы деревьев решений. Тема 5.6 Ансамбли деревьев решений: Случайный лес. Тема 5.7 Ансамбли деревьев решений: Градиентный бустинг.</p>	<p>Обсуждение решений профессионально-ориентированных заданий и задач; обсуждение и анализ решения кейсов</p>	16

**РАЗДЕЛ 6. УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ
САМОСТОЯТЕЛЬНОЙ РАБОТЫ ОБУЧАЮЩИХСЯ ПО ДИСЦИПЛИНЕ**

Самостоятельная работа

Наименование разделов/тем	Вопросы, выносимые на самостоятельное изучение
<p>Раздел №1 «Введение в статистические методы анализа данных» Тема 1.1 Введение в статистические методы анализа данных. Тема 1.2 Методы первичной обработки и представления данных. Тема 1.3 Методы выявления и устранения аномальных наблюдений.</p>	<ol style="list-style-type: none"> 1. Информационный подход к анализу данных. 2. Приложение методов обработки данных к решению практических задач. 3. Методы выявления аномальных наблюдений. 4. Методы устранения аномальных наблюдений.
<p>Раздел №2 «Множественный корреляционно-</p>	<ol style="list-style-type: none"> 1. Этапы построения модели множествен-

Наименование разделов/тем	Вопросы, выносимые на самостоятельное изучение
<p>регрессионный анализ. Классическая модель множественной регрессии» Тема 2.1 Корреляционно-регрессионный анализ. Модель множественной регрессии. Тема 2.2 Фиктивные переменные. Критерий Чоу проверки однородности. Тема 2.3 Множественные коэффициенты корреляции. Тема 2.4 Статистический анализ адекватности модели множественной регрессии. Тема 2.5 Мультиколлинеарность факторов.</p>	<p>ной регрессии. 2. Методы формирования выборки наблюдений фиктивных переменных для построения регрессии. 3. Коэффициент детерминации. Свойства. 4. Доверительные интервалы для теоретических коэффициентов регрессии. 5. Практическая реализация проверки адекватности модели множественной регрессии. 6. Мультиколлинеарность.</p>
<p>Раздел №3 «Обобщенная линейная модель множественной регрессии. Регрессионные модели с автокоррелированными остатками» Тема 3.1 Обобщенная линейная модель множественной регрессии. Обобщенный МНК Тема 3.2 Гетероскедастичность. Методы выявления гетероскедастичности. Тема 3.3 Взвешенный МНК. Устранение гетероскедастичности. Тема 3.4 Регрессионные модели с автокоррелированными остатками. Тема 3.5 Устранение автокорреляции</p>	<p>1. Взвешенный МНК. 2. Тесты устранения гетероскедастичности. 3. Примеры устранения гетероскедастичности. 4. Примеры практической реализации критериев проверки автокорреляции остатков. 5. Примеры практической реализации критериев проверки автокорреляции остатков. 6. Методы построения оценки коэффициента автокорреляции первого порядка.</p>
<p>Раздел №4 «Кластерный анализ» Тема 4.1 Кластерный анализ. Основные этапы кластерного анализа. Тема 4.2 Метрики расстояний. Свойства кластеров. Методы кластерного анализа. Тема 4.3 Расстояния между кластерами. Методы группировки кластеров. Тема 4.4 Иерархические методы кластеризации. Агломеративные методы. Тема 4.5 Иерархические дивизимные методы кластеризации. Тема 4.6 Примеры реализации иерархических методов кластеризации. Тема 4.7 Итеративные методы. Метод к-средних. Тема 4.8 Метод к-средних. Критерии качества кластеризации. Тема 4.9 Метод нечетких к-средних</p>	<p>1. Основные задачи и приложения кластерного анализа. 2. Классификация методов кластерного анализа. 3. Дисперсионные методы группировки кластеров. 4. Иерархические методы кластеризации. Практическое приложение. 5. Дендограмма. 6. Пример реализации метода к-средних. 7. Статистические критерии качества разбиения. 8. Пример реализации метода нечетких к-средних</p>
<p>Раздел №5 «Методы классификации» Тема 5.1 Понятие методов классификации. Тема 5.2 Логистическая регрессия. Тема 5.3 Метрики качества классификации. Тема 5.4 Метрики качества классификации. Регуляризация в задаче регрессии. Тема 5.5 Методы деревьев решений. Тема 5.6 Ансамбли деревьев решений: Случайный лес. Тема 5.7 Ансамбли деревьев решений: Градиентный бустинг.</p>	<p>1. Примеры задач классификации. 2. Методы оценки параметров логистической регрессии. 3. Построение ROC-кривой. 4. Основные алгоритмы построения деревьев решений. 5. Пример практической реализации алгоритма случайного леса. 6. Пример практической реализации алгоритма градиентного бустинга.</p>

6.1. Примерные задания для самостоятельной работы

1. Для предложенной выборки наблюдений провести проверку наличия аномальных наблюдений, сравнить результаты реализации критериев.

2. Построить эмпирическое уравнение множественной линейной регрессии для выбранного набора зависимой и объясняющих переменных, провести проверку адекватности построенного уравнения регрессии, вычислить основные метрики качества регрессии.
3. Для предложенной выборки наблюдений требуется построить оптимальное разбиение субъектов, используя метод к-средних, нечетких к-средних.
4. Для предложенной выборки наблюдений необходимо построить модели классификации при различных параметрах моделирования. На основе проведенных расчетов провести проверку качества построенных моделей с использованием метрик, выбрать оптимальную модель классификации, обосновать свой выбор.

РАЗДЕЛ 7. ОЦЕНОЧНЫЕ И МЕТОДИЧЕСКИЕ МАТЕРИАЛЫ ПО ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЕ (ФОНД ОЦЕНОЧНЫХ СРЕДСТВ) ДЛЯ ПРОВЕДЕНИЯ ТЕКУЩЕГО КОНТРОЛЯ УСПЕВАЕМОСТИ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ

7.1. Форма промежуточной аттестации обучающегося по учебной дисциплине.

Контрольным мероприятием промежуточной аттестации обучающихся по учебной дисциплине «Статистические методы анализа данных» в 5 является зачет с оценкой, в 6 семестрах является экзамен, который проводится в устной форме.

Таблица 7.1

ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ, СООТНЕСЕННЫЕ С ПЛАНИРУЕМЫМИ РЕЗУЛЬТАТАМИ ОСВОЕНИЯ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ В СООТНОШЕНИИ С ОЦЕНОЧНЫМИ СРЕДСТВАМИ

Планируемые результаты, характеризующие этапы формирования компетенции	Содержание учебного материала	Примеры контрольных вопросов и заданий для оценки знаний, умений, владений	Методы/ средства контроля
УК-1 Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач			
ИУК-1.1 Знать методы поиска информации для решения поставленной задачи	<p>Раздел 1 «Введение в статистические методы анализа данных»</p> <p>Тема 1.1 Введение в статистические методы анализа данных.</p> <p>Тема 1.2 Методы первичной обработки и представления данных.</p> <p>Тема 1.3 Методы выявления и устранения аномальных наблюдений.</p> <p>Раздел №2 «Множественный корреляционно-регрессионный анализ. Классическая модель множественной регрессии»</p> <p>Тема 2.1 Корреляционно-регрессионный анализ. Модель множественной регрессии.</p> <p>Тема 2.2 Фиктивные переменные. Критерий Чоу проверки однородности.</p> <p>Тема 2.3 Множественные коэффициенты корреляции.</p> <p>Тема 2.4 Статистический анализ адекватности модели множественной регрессии.</p> <p>Тема 2.5 Мультиколлинеарность факторов.</p> <p>Раздел №3 «Обобщенная линейная модель множественной регрессии. Регрессионные модели с автокоррелированными остатками»</p> <p>Тема 3.1 Обобщенная линейная</p>	<ol style="list-style-type: none"> 1. Основные этапы эмпирического исследования. 2. Дискретные, непрерывные данные. Пространственная выборка, временной ряд, панельные данные. 3. Аномальные наблюдения. Критерии проверки аномальности наблюдений. 4. Общий вид модели множественной регрессии. Линейная модель множественной регрессии. Матричная форма. 5. Основные предпосылки регрессионного анализа. Метод наименьших квадратов. МНК-оценки параметров множественной линейной регрессии. Теорема Гаусса-Маркова. 6. Статистические свойства МНК оценок. Проверка гипотезы о статистической значимости теоретических коэффициентов регрессии. Доверительные интервалы для теоретических коэффициентов. 7. Частный коэффициент корреляции, множественный коэффициент корреляции. Проверка гипотезы о статистической значимости частного коэффициента корреляции. 8. Коэффициент детерминации. Свойства. Дисперсион- 	<p>Устный контроль/ опрос на семинарских занятиях, зачете, экзамене; анализ докладов на семинарских занятиях; анализ защиты рефератов; анализ защиты проектов; применение теоретических знаний при анализе (разборе) конкретных практико-ориентированных ситуаций и профессионально-прикладных задач, анализ использования теоретических знаний в процессе решения кейсов, в ходе деловых игр; письменный контроль, анализ содержания эссе; тестирование (выполнение тестовых заданий)</p>

	<p>модель множественной регрессии. Обобщенный МНК Тема 3.2 Гетероскедастичность. Методы выявления гетероскедастичности. Тема 3.3 Взвешенный МНК. Устранение гетероскедастичности. Тема 3.4 Регрессионные модели с автокоррелированными остатками. Тема 3.5 Устранение автокорреляции Раздел №4 «Кластерный анализ» Тема 4.1 Кластерный анализ. Основные этапы кластерного анализа. Тема 4.2 Метрики расстояний. Свойства кластеров. Методы кластерного анализа. Тема 4.3 Расстояния между кластерами. Методы группировки кластеров. Тема 4.4 Иерархические методы кластеризации. Агломеративные методы. Тема 4.5 Иерархические дивизимные методы кластеризации. Тема 4.6 Примеры реализации иерархических методов кластеризации. Тема 4.7 Итеративные методы. Метод k-средних. Тема 4.8 Метод k-средних. Критерии качества кластеризации. Тема 4.9 Метод нечетких k-средних Раздел №5 «Методы классификации» Тема 5.1 Понятие методов классификации. Тема 5.2 Логистическая регрессия. Тема 5.3 Метрики качества классификации. Тема 5.4 Метрики качества классификации. Регуляризация в задаче регрессии. Тема 5.5 Методы деревьев решений. Тема 5.6 Ансамбли деревьев решений: Случайный лес. Тема 5.7 Ансамбли деревьев решений: Градиентный бустинг.</p>	<p>ное отношение Фишера. Проверка гипотезы о статистической значимости уравнения регрессии. 9. Обобщенная линейная модель множественной регрессии. Обобщенный МНК. 10. Гомоскедастичность, гетероскедастичность. Критерии проверки остатков на гетероскедастичность. Методы устранения гетероскедастичности. 11. Автокорреляция остатков. Критерии проверки автокорреляции. Устранение автокорреляции. 12. Понятие кластерного анализа. Основные этапы кластерного анализа. 13. Основные меры сходства объектов. Расстояние между объектами: основные виды, формулы расчета. 14. Свойства кластера: плотность, дисперсия, размеры, форма, отделимость. Расстояние между кластерами: основные виды, формулы расчета. 15. Иерархические агломеративные методы: понятие, алгоритм реализации. 16. Иерархические дивизимные методы: понятие, алгоритм реализации. 17. Метод k-средних: понятие, алгоритм реализации. Функционалы качества разбиения объектов методом k-средних. 18. Метод нечетких k-средних: понятие, алгоритм реализации. Функционалы качества разбиения объектов методом k-средних. 19. Постановка задачи классификации. Основные методы классификации. 20. Общий вид логистической регрессии. Оценка логистической регрессии. Понятие сбалансированной выборки. 21. Метрики качества классификации. 22. Методы деревьев решений. 23. Понятие ансамблей решений. Случайный лес. Градиентный бустинг.</p>	
<p>ИУК-1.2 Уметь проводить критический анализ и синтез информации</p>	<p>Раздел 1 «Введение в статистические методы анализа данных» Тема 1.1 Введение в статистические методы анализа данных. Тема 1.2 Методы первичной обработки и представления данных. Тема 1.3 Методы выявления и устранения аномальных наблюдений. Раздел №2 «Множественный корреляционно-регрессионный</p>	<p>1. Основные этапы эмпирического исследования. 2. Дискретные, непрерывные данные. Пространственная выборка, временной ряд, панельные данные. 3. Аномальные наблюдения. Критерии проверки аномальности наблюдений. 4. Общий вид модели множественной регрессии. Линейная модель множественной</p>	<p>Анализ проявленных умений при решении кейсов, в ходе деловых игр; письменный контроль, анализ качества решений профессиональных задач в контрольных работах; анализ содержания</p>

	<p>анализ. Классическая модель множественной регрессии» Тема 2.1 Корреляционно-регрессионный анализ. Модель множественной регрессии. Тема 2.2 Фиктивные переменные. Критерий Чоу проверки однородности. Тема 2.3 Множественные коэффициенты корреляции. Тема 2.4 Статистический анализ адекватности модели множественной регрессии. Тема 2.5 Мультиколлинеарность факторов. Раздел №3 «Обобщенная линейная модель множественной регрессии. Регрессионные модели с автокоррелированными остатками» Тема 3.1 Обобщенная линейная модель множественной регрессии. Обобщенный МНК Тема 3.2 Гетероскедастичность. Методы выявления гетероскедастичности. Тема 3.3 Взвешенный МНК. Устранение гетероскедастичности. Тема 3.4 Регрессионные модели с автокоррелированными остатками. Тема 3.5 Устранение автокорреляции Раздел №4 «Кластерный анализ» Тема 4.1 Кластерный анализ. Основные этапы кластерного анализа. Тема 4.2 Метрики расстояний. Свойства кластеров. Методы кластерного анализа. Тема 4.3 Расстояния между кластерами. Методы группировки кластеров. Тема 4.4 Иерархические методы кластеризации. Агломеративные методы. Тема 4.5 Иерархические дивизимные методы кластеризации. Тема 4.6 Примеры реализации иерархических методов кластеризации. Тема 4.7 Итеративные методы. Метод k-средних. Тема 4.8 Метод k-средних. Критерии качества кластеризации. Тема 4.9 Метод нечетких k-средних Раздел №5 «Методы классификации» Тема 5.1 Понятие методов классификации. Тема 5.2 Логистическая регрессия. Тема 5.3 Метрики качества клас-</p>	<p>регрессии. Матричная форма. 5. Основные предпосылки регрессионного анализа. Метод наименьших квадратов. МНК-оценки параметров множественной линейной регрессии. Теорема Гаусса-Маркова. 6. Статистические свойства МНК оценок. Проверка гипотезы о статистической значимости теоретических коэффициентов регрессии. Доверительные интервалы для теоретических коэффициентов. 7. Частный коэффициент корреляции, множественный коэффициент корреляции. Проверка гипотезы о статистической значимости частного коэффициента корреляции. 8. Коэффициент детерминации. Свойства. Дисперсионное отношение Фишера. Проверка гипотезы о статистической значимости уравнения регрессии. 9. Обобщенная линейная модель множественной регрессии. Обобщенный МНК. 10. Гомоскедастичность, гетероскедастичность. Критерии проверки остатков на гетероскедастичность. Методы устранения гетероскедастичности. 11. Автокорреляция остатков. Критерии проверки автокорреляции. Устранение автокорреляции. 12. Понятие кластерного анализа. Основные этапы кластерного анализа. 13. Основные меры сходства объектов. Расстояние между объектами: основные виды, формулы расчета. 14. Свойства кластера: плотность, дисперсия, размеры, форма, отделимость. Расстояние между кластерами: основные виды, формулы расчета. 15. Иерархические агломеративные методы: понятие, алгоритм реализации. 16. Иерархические дивизимные методы: понятие, алгоритм реализации. 17. Метод k-средних: понятие, алгоритм реализации. Функционалы качества разбиения объектов методом k-средних. 18. Метод нечетких k-средних: понятие, алгоритм реализации. Функционалы качества разбиения объектов методом k-средних. 19. Постановка задачи классификации. Основные методы классификации.</p>	<p>профессионально-ориентированных эссе; тестирование (выполнение тестовых заданий); анализ защит профессионально-ориентированных проектов; опрос на семинарских занятиях, зачете, анализ докладов на семинарских занятиях; анализ защиты рефератов; анализ решения конкретных практико-ориентированных ситуаций и профессионально-прикладных задач, анализ выполнения контрольных работ</p>
--	---	--	--

	<p>сификации. Тема 5.4 Метрики качества классификации. Регуляризация в задаче регрессии. Тема 5.5 Методы деревьев решений. Тема 5.6 Ансамбли деревьев решений: Случайный лес. Тема 5.7 Ансамбли деревьев решений: Градиентный бустинг.</p>	<p>20. Общий вид логистической регрессии. Оценка логистической регрессии. Понятие сбалансированной выборки. 21. Метрики качества классификации. 22. Методы деревьев решений. 23. Понятие ансамблей решений. Случайный лес. Градиентный бустинг.</p>	
<p>ИУК-1.3 Владеть системным подходом для решения поставленных задач</p>	<p>Раздел 1 «Введение в статистические методы анализа данных» Тема 1.1 Введение в статистические методы анализа данных. Тема 1.2 Методы первичной обработки и представления данных. Тема 1.3 Методы выявления и устранения аномальных наблюдений. Раздел №2 «Множественный корреляционно-регрессионный анализ. Классическая модель множественной регрессии» Тема 2.1 Корреляционно-регрессионный анализ. Модель множественной регрессии. Тема 2.2 Фиктивные переменные. Критерий Чоу проверки однородности. Тема 2.3 Множественные коэффициенты корреляции. Тема 2.4 Статистический анализ адекватности модели множественной регрессии. Тема 2.5 Мультиколлинеарность факторов. Раздел №3 «Обобщенная линейная модель множественной регрессии. Регрессионные модели с автокоррелированными остатками» Тема 3.1 Обобщенная линейная модель множественной регрессии. Обобщенный МНК Тема 3.2 Гетероскедастичность. Методы выявления гетероскедастичности. Тема 3.3 Взвешенный МНК. Устранение гетероскедастичности. Тема 3.4 Регрессионные модели с автокоррелированными остатками. Тема 3.5 Устранение автокорреляции Раздел №4 «Кластерный анализ» Тема 4.1 Кластерный анализ. Основные этапы кластерного анализа. Тема 4.2 Метрики расстояний. Свойства кластеров. Методы кластерного анализа. Тема 4.3 Расстояния между кластерами. Методы группировки кластеров.</p>	<p>1. Основные этапы эмпирического исследования. 2. Дискретные, непрерывные данные. Пространственная выборка, временной ряд, панельные данные. 3. Аномальные наблюдения. Критерии проверки аномальности наблюдений. 4. Общий вид модели множественной регрессии. Линейная модель множественной регрессии. Матричная форма. 5. Основные предпосылки регрессионного анализа. Метод наименьших квадратов. МНК-оценки параметров множественной линейной регрессии. Теорема Гаусса-Маркова. 6. Статистические свойства МНК оценок. Проверка гипотезы о статистической значимости теоретических коэффициентов регрессии. Доверительные интервалы для теоретических коэффициентов. 7. Частный коэффициент корреляции, множественный коэффициент корреляции. Проверка гипотезы о статистической значимости частного коэффициента корреляции. 8. Коэффициент детерминации. Свойства. Дисперсионное отношение Фишера. Проверка гипотезы о статистической значимости уравнения регрессии. 9. Обобщенная линейная модель множественной регрессии. Обобщенный МНК. 10. Гомоскедастичность, гетероскедастичность. Критерии проверки остатков на гетероскедастичность. Методы устранения гетероскедастичности. 11. Автокорреляция остатков. Критерии проверки автокорреляции. Устранение автокорреляции. 12. Понятие кластерного анализа. Основные этапы кластерного анализа. 13. Основные меры сходства объектов. Расстояние между объектами: основные виды, формулы расчета. 14. Свойства кластера:</p>	<p>Анализ проявленных навыков при решении кейсов, в ходе деловых игр; письменный контроль, анализ качества решений профессиональных задач в контрольных работах; анализ содержания профессионально-ориентированных эссе; тестирование (выполнение тестовых заданий); анализ защит профессионально-ориентированных проектов; опрос на семинарских занятиях, зачете, экзамене; анализ докладов на семинарских занятиях; анализ защиты рефератов; анализ решения конкретных практико-ориентированных ситуаций и профессионально-прикладных задач, анализ выполнения контрольных работ</p>

	<p>Тема 4.4 Иерархические методы кластеризации. Агломеративные методы.</p> <p>Тема 4.5 Иерархические дивизимные методы кластеризации.</p> <p>Тема 4.6 Примеры реализации иерархических методов кластеризации.</p> <p>Тема 4.7 Итеративные методы. Метод k-средних.</p> <p>Тема 4.8 Метод k-средних. Критерии качества кластеризации.</p> <p>Тема 4.9 Метод нечетких k-средних</p> <p>Раздел №5 «Методы классификации»</p> <p>Тема 5.1 Понятие методов классификации.</p> <p>Тема 5.2 Логистическая регрессия.</p> <p>Тема 5.3 Метрики качества классификации.</p> <p>Тема 5.4 Метрики качества классификации. Регуляризация в задаче регрессии.</p> <p>Тема 5.5 Методы деревьев решений.</p> <p>Тема 5.6 Ансамбли деревьев решений: Случайный лес.</p> <p>Тема 5.7 Ансамбли деревьев решений: Градиентный бустинг.</p>	<p>плотность, дисперсия, размеры, форма, отделимость. Расстояние между кластерами: основные виды, формулы расчета.</p> <p>15. Иерархические агломеративные методы: понятие, алгоритм реализации.</p> <p>16. Иерархические дивизимные методы: понятие, алгоритм реализации.</p> <p>17. Метод k-средних: понятие, алгоритм реализации. Функционалы качества разбиения объектов методом k-средних.</p> <p>18. Метод нечетких k-средних: понятие, алгоритм реализации. Функционалы качества разбиения объектов методом k-средних.</p> <p>19. Постановка задачи классификации. Основные методы классификации.</p> <p>20. Общий вид логистической регрессии. Оценка логистической регрессии. Понятие сбалансированной выборки.</p> <p>21. Метрики качества классификации.</p> <p>22. Методы деревьев решений.</p> <p>23. Понятие ансамблей решений. Случайный лес. Градиентный бустинг.</p>	
ПК-1 Способность проектировать прикладное программное обеспечение			
<p>ИПК-1.1 Знать принципы построения архитектуры программного обеспечения и виды архитектуры программного обеспечения; типовые решения, библиотеки программных модулей, шаблоны, классы объектов, используемые при разработке программного обеспечения; методы и средства проектирования программного обеспечения; методы и средства проектирования баз данных; методы и средства проектирования программных интерфейсов</p>	<p>Раздел 1 «Введение в статистические методы анализа данных»</p> <p>Тема 1.1 Введение в статистические методы анализа данных.</p> <p>Тема 1.2 Методы первичной обработки и представления данных.</p> <p>Тема 1.3 Методы выявления и устранения аномальных наблюдений.</p> <p>Раздел №2 «Множественный корреляционно-регрессионный анализ. Классическая модель множественной регрессии»</p> <p>Тема 2.1 Корреляционно-регрессионный анализ. Модель множественной регрессии.</p> <p>Тема 2.2 Фиктивные переменные. Критерий Чоу проверки однородности.</p> <p>Тема 2.3 Множественные коэффициенты корреляции.</p> <p>Тема 2.4 Статистический анализ адекватности модели множественной регрессии.</p> <p>Тема 2.5 Мультиколлинеарность факторов.</p> <p>Раздел №3 «Обобщенная линейная модель множественной регрессии. Регрессионные модели с автокоррелированными остатками»</p> <p>Тема 3.1 Обобщенная линейная</p>	<ol style="list-style-type: none"> 1. Основные этапы эмпирического исследования. 2. Дискретные, непрерывные данные. Пространственная выборка, временной ряд, панельные данные. 3. Аномальные наблюдения. Критерии проверки аномальности наблюдений. 4. Общий вид модели множественной регрессии. Линейная модель множественной регрессии. Матричная форма. 5. Основные предпосылки регрессионного анализа. Метод наименьших квадратов. МНК-оценки параметров множественной линейной регрессии. Теорема Гаусса-Маркова. 6. Статистические свойства МНК оценок. Проверка гипотезы о статистической значимости теоретических коэффициентов регрессии. Доверительные интервалы для теоретических коэффициентов. 7. Частный коэффициент корреляции, множественный коэффициент корреляции. Проверка гипотезы о статистической значимости частного коэффициента корреляции. 8. Коэффициент детерминации. Свойства. Дисперсион- 	<p>Устный контроль/ опрос на семинарских занятиях, зачете, экзамене; анализ докладов на семинарских занятиях; анализ защиты рефератов; анализ защиты проектов; применение теоретических знаний при анализе (разборе) конкретных практико-ориентированных ситуаций и профессионально-прикладных задач, анализ использования теоретических знаний в процессе решения кейсов, в ходе деловых игр; письменный контроль, анализ содержания эссе; тестирование (выполнение тестовых заданий)</p>

	<p>модель множественной регрессии. Обобщенный МНК Тема 3.2 Гетероскедастичность. Методы выявления гетероскедастичности. Тема 3.3 Взвешенный МНК. Устранение гетероскедастичности. Тема 3.4 Регрессионные модели с автокоррелированными остатками. Тема 3.5 Устранение автокорреляции Раздел №4 «Кластерный анализ» Тема 4.1 Кластерный анализ. Основные этапы кластерного анализа. Тема 4.2 Метрики расстояний. Свойства кластеров. Методы кластерного анализа. Тема 4.3 Расстояния между кластерами. Методы группировки кластеров. Тема 4.4 Иерархические методы кластеризации. Агломеративные методы. Тема 4.5 Иерархические дивизимные методы кластеризации. Тема 4.6 Примеры реализации иерархических методов кластеризации. Тема 4.7 Итеративные методы. Метод k-средних. Тема 4.8 Метод k-средних. Критерии качества кластеризации. Тема 4.9 Метод нечетких k-средних Раздел №5 «Методы классификации» Тема 5.1 Понятие методов классификации. Тема 5.2 Логистическая регрессия. Тема 5.3 Метрики качества классификации. Тема 5.4 Метрики качества классификации. Регуляризация в задаче регрессии. Тема 5.5 Методы деревьев решений. Тема 5.6 Ансамбли деревьев решений: Случайный лес. Тема 5.7 Ансамбли деревьев решений: Градиентный бустинг.</p>	<p>ное отношение Фишера. Проверка гипотезы о статистической значимости уравнения регрессии. 9. Обобщенная линейная модель множественной регрессии. Обобщенный МНК. 10. Гомоскедастичность, гетероскедастичность. Критерии проверки остатков на гетероскедастичность. Методы устранения гетероскедастичности. 11. Автокорреляция остатков. Критерии проверки автокорреляции. Устранение автокорреляции. 12. Понятие кластерного анализа. Основные этапы кластерного анализа. 13. Основные меры сходства объектов. Расстояние между объектами: основные виды, формулы расчета. 14. Свойства кластера: плотность, дисперсия, размеры, форма, отделимость. Расстояние между кластерами: основные виды, формулы расчета. 15. Иерархические агломеративные методы: понятие, алгоритм реализации. 16. Иерархические дивизимные методы: понятие, алгоритм реализации. 17. Метод k-средних: понятие, алгоритм реализации. Функционалы качества разбиения объектов методом k-средних. 18. Метод нечетких k-средних: понятие, алгоритм реализации. Функционалы качества разбиения объектов методом k-средних. 19. Постановка задачи классификации. Основные методы классификации. 20. Общий вид логистической регрессии. Оценка логистической регрессии. Понятие сбалансированной выборки. 21. Метрики качества классификации. 22. Методы деревьев решений. 23. Понятие ансамблей решений. Случайный лес. Градиентный бустинг.</p>	
<p>ИПК-1.2 Уметь использовать существующие типовые решения и шаблоны проектирования программного обеспечения; применять методы и средства проектирования программного обеспечения, структур данных, баз данных, программных интерфейсов; осуществлять коммуникации с заинтересованными сторо-</p>	<p>Раздел 1 «Введение в статистические методы анализа данных» Тема 1.1 Введение в статистические методы анализа данных. Тема 1.2 Методы первичной обработки и представления данных. Тема 1.3 Методы выявления и устранения аномальных наблюдений. Раздел №2 «Множественный корреляционно-регрессионный</p>	<p>1. Основные этапы эмпирического исследования. 2. Дискретные, непрерывные данные. Пространственная выборка, временной ряд, панельные данные. 3. Аномальные наблюдения. Критерии проверки аномальности наблюдений. 4. Общий вид модели множественной регрессии. Линейная модель множественной</p>	<p>Анализ проявленных умений при решении кейсов, в ходе деловых игр; письменный контроль, анализ качества решений профессиональных задач в контрольных работах; анализ содержания</p>

нами	<p>анализ. Классическая модель множественной регрессии» Тема 2.1 Корреляционно-регрессионный анализ. Модель множественной регрессии. Тема 2.2 Фиктивные переменные. Критерий Чоу проверки однородности. Тема 2.3 Множественные коэффициенты корреляции. Тема 2.4 Статистический анализ адекватности модели множественной регрессии. Тема 2.5 Мультиколлинеарность факторов. Раздел №3 «Обобщенная линейная модель множественной регрессии. Регрессионные модели с автокоррелированными остатками» Тема 3.1 Обобщенная линейная модель множественной регрессии. Обобщенный МНК Тема 3.2 Гетероскедастичность. Методы выявления гетероскедастичности. Тема 3.3 Взвешенный МНК. Устранение гетероскедастичности. Тема 3.4 Регрессионные модели с автокоррелированными остатками. Тема 3.5 Устранение автокорреляции Раздел №4 «Кластерный анализ» Тема 4.1 Кластерный анализ. Основные этапы кластерного анализа. Тема 4.2 Метрики расстояний. Свойства кластеров. Методы кластерного анализа. Тема 4.3 Расстояния между кластерами. Методы группировки кластеров. Тема 4.4 Иерархические методы кластеризации. Агломеративные методы. Тема 4.5 Иерархические дивизимные методы кластеризации. Тема 4.6 Примеры реализации иерархических методов кластеризации. Тема 4.7 Итеративные методы. Метод k-средних. Тема 4.8 Метод k-средних. Критерии качества кластеризации. Тема 4.9 Метод нечетких k-средних Раздел №5 «Методы классификации» Тема 5.1 Понятие методов классификации. Тема 5.2 Логистическая регрессия. Тема 5.3 Метрики качества клас-</p>	<p>регрессии. Матричная форма. 5. Основные предпосылки регрессионного анализа. Метод наименьших квадратов. МНК-оценки параметров множественной линейной регрессии. Теорема Гаусса-Маркова. 6. Статистические свойства МНК оценок. Проверка гипотезы о статистической значимости теоретических коэффициентов регрессии. Доверительные интервалы для теоретических коэффициентов. 7. Частный коэффициент корреляции, множественный коэффициент корреляции. Проверка гипотезы о статистической значимости частного коэффициента корреляции. 8. Коэффициент детерминации. Свойства. Дисперсионное отношение Фишера. Проверка гипотезы о статистической значимости уравнения регрессии. 9. Обобщенная линейная модель множественной регрессии. Обобщенный МНК. 10. Гомоскедастичность, гетероскедастичность. Критерии проверки остатков на гетероскедастичность. Методы устранения гетероскедастичности. 11. Автокорреляция остатков. Критерии проверки автокорреляции. Устранение автокорреляции. 12. Понятие кластерного анализа. Основные этапы кластерного анализа. 13. Основные меры сходства объектов. Расстояние между объектами: основные виды, формулы расчета. 14. Свойства кластера: плотность, дисперсия, размеры, форма, отделимость. Расстояние между кластерами: основные виды, формулы расчета. 15. Иерархические агломеративные методы: понятие, алгоритм реализации. 16. Иерархические дивизимные методы: понятие, алгоритм реализации. 17. Метод k-средних: понятие, алгоритм реализации. Функционалы качества разбиения объектов методом k-средних. 18. Метод нечетких k-средних: понятие, алгоритм реализации. Функционалы качества разбиения объектов методом k-средних. 19. Постановка задачи классификации. Основные методы классификации.</p>	<p>профессионально-ориентированных эссе; тестирование (выполнение тестовых заданий); анализ защит профессионально-ориентированных проектов; опрос на семинарских занятиях, зачете, анализ докладов на семинарских занятиях; анализ защиты рефератов; анализ решения конкретных практико-ориентированных ситуаций и профессионально-прикладных задач, анализ выполнения контрольных работ</p>
------	---	--	--

	<p>сификации. Тема 5.4 Метрики качества классификации. Регуляризация в задаче регрессии. Тема 5.5 Методы деревьев решений. Тема 5.6 Ансамбли деревьев решений: Случайный лес. Тема 5.7 Ансамбли деревьев решений: Градиентный бустинг.</p>	<p>20. Общий вид логистической регрессии. Оценка логистической регрессии. Понятие сбалансированной выборки. 21. Метрики качества классификации. 22. Методы деревьев решений. 23. Понятие ансамблей решений. Случайный лес. Градиентный бустинг.</p>	
<p>ИПК-1.3 Владеть разработкой, изменением и согласованием архитектуры программного обеспечения с системным аналитиком и архитектором программного обеспечения; системой проектирования структур данных; системой проектирования баз данных; системой проектирования программных интерфейсов; оценкой и согласованием сроков выполнения поставленных задач</p>	<p>Раздел 1 «Введение в статистические методы анализа данных» Тема 1.1 Введение в статистические методы анализа данных. Тема 1.2 Методы первичной обработки и представления данных. Тема 1.3 Методы выявления и устранения аномальных наблюдений. Раздел №2 «Множественный корреляционно-регрессионный анализ. Классическая модель множественной регрессии» Тема 2.1 Корреляционно-регрессионный анализ. Модель множественной регрессии. Тема 2.2 Фиктивные переменные. Критерий Чоу проверки однородности. Тема 2.3 Множественные коэффициенты корреляции. Тема 2.4 Статистический анализ адекватности модели множественной регрессии. Тема 2.5 Мультиколлинеарность факторов. Раздел №3 «Обобщенная линейная модель множественной регрессии. Регрессионные модели с автокоррелированными остатками» Тема 3.1 Обобщенная линейная модель множественной регрессии. Обобщенный МНК Тема 3.2 Гетероскедастичность. Методы выявления гетероскедастичности. Тема 3.3 Взвешенный МНК. Устранение гетероскедастичности. Тема 3.4 Регрессионные модели с автокоррелированными остатками. Тема 3.5 Устранение автокорреляции Раздел №4 «Кластерный анализ» Тема 4.1 Кластерный анализ. Основные этапы кластерного анализа. Тема 4.2 Метрики расстояний. Свойства кластеров. Методы кластерного анализа. Тема 4.3 Расстояния между кластерами. Методы группировки кластеров.</p>	<p>1. Основные этапы эмпирического исследования. 2. Дискретные, непрерывные данные. Пространственная выборка, временной ряд, панельные данные. 3. Аномальные наблюдения. Критерии проверки аномальности наблюдений. 4. Общий вид модели множественной регрессии. Линейная модель множественной регрессии. Матричная форма. 5. Основные предпосылки регрессионного анализа. Метод наименьших квадратов. МНК-оценки параметров множественной линейной регрессии. Теорема Гаусса-Маркова. 6. Статистические свойства МНК оценок. Проверка гипотезы о статистической значимости теоретических коэффициентов регрессии. Доверительные интервалы для теоретических коэффициентов. 7. Частный коэффициент корреляции, множественный коэффициент корреляции. Проверка гипотезы о статистической значимости частного коэффициента корреляции. 8. Коэффициент детерминации. Свойства. Дисперсионное отношение Фишера. Проверка гипотезы о статистической значимости уравнения регрессии. 9. Обобщенная линейная модель множественной регрессии. Обобщенный МНК. 10. Гомоскедастичность, гетероскедастичность. Критерии проверки остатков на гетероскедастичность. Методы устранения гетероскедастичности. 11. Автокорреляция остатков. Критерии проверки автокорреляции. Устранение автокорреляции. 12. Понятие кластерного анализа. Основные этапы кластерного анализа. 13. Основные меры сходства объектов. Расстояние между объектами: основные виды, формулы расчета. 14. Свойства кластера:</p>	<p>Анализ проявленных навыков при решении кейсов, в ходе деловых игр; письменный контроль, анализ качества решений профессиональных задач в контрольных работах; анализ содержания профессионально-ориентированных эссе; тестирование (выполнение тестовых заданий); анализ защит профессионально-ориентированных проектов; опрос на семинарских занятиях, зачете, экзамене; анализ докладов на семинарских занятиях; анализ защиты рефератов; анализ решения конкретных практико-ориентированных ситуаций и профессионально-прикладных задач, анализ выполнения контрольных работ</p>

	<p>Тема 4.4 Иерархические методы кластеризации. Агломеративные методы.</p> <p>Тема 4.5 Иерархические дивизимные методы кластеризации.</p> <p>Тема 4.6 Примеры реализации иерархических методов кластеризации.</p> <p>Тема 4.7 Итеративные методы. Метод к-средних.</p> <p>Тема 4.8 Метод к-средних. Критерии качества кластеризации.</p> <p>Тема 4.9 Метод нечетких к-средних</p> <p>Раздел №5 «Методы классификации»</p> <p>Тема 5.1 Понятие методов классификации.</p> <p>Тема 5.2 Логистическая регрессия.</p> <p>Тема 5.3 Метрики качества классификации.</p> <p>Тема 5.4 Метрики качества классификации. Регуляризация в задаче регрессии.</p> <p>Тема 5.5 Методы деревьев решений.</p> <p>Тема 5.6 Ансамбли деревьев решений: Случайный лес.</p> <p>Тема 5.7 Ансамбли деревьев решений: Градиентный бустинг.</p>	<p>плотность, дисперсия, размеры, форма, отделимость. Расстояние между кластерами: основные виды, формулы расчета.</p> <p>15. Иерархические агломеративные методы: понятие, алгоритм реализации.</p> <p>16. Иерархические дивизимные методы: понятие, алгоритм реализации.</p> <p>17. Метод к-средних: понятие, алгоритм реализации. Функционалы качества разбиения объектов методом к-средних.</p> <p>18. Метод нечетких к-средних: понятие, алгоритм реализации. Функционалы качества разбиения объектов методом к-средних.</p> <p>19. Постановка задачи классификации. Основные методы классификации.</p> <p>20. Общий вид логистической регрессии. Оценка логистической регрессии. Понятие сбалансированной выборки.</p> <p>21. Метрики качества классификации.</p> <p>22. Методы деревьев решений.</p> <p>23. Понятие ансамблей решений. Случайный лес. Градиентный бустинг.</p>	
<p>ПК-2. Способность осуществлять инженерно-технологическую поддержку планирования управления требованиями по созданию (модификации) и сопровождению ИС</p>			
<p>ИПК-2.1. Знать инструменты и методы управления требованиями; предметной областью автоматизации; возможностями ИС; источниками информации, необходимой для профессиональной деятельности; современный отечественный и зарубежный опыт в профессиональной деятельности; инструменты и методы моделирования бизнес-процессов организации; основы реинжиниринга бизнес-процессов организации; управление содержанием проекта: документирование требований, систему анализа продукта, модернизацию совещаний; управление качеством: контрольные списки, верификация, валидация (приемо-сдаточные испытания); управление коммуникациями в проекте: базовые навыки управления (в том числе проведение презентаций, проведение переговоров, публичные выступления); культуру речи; правила деловой переписки.</p>	<p>Раздел 1 «Введение в статистические методы анализа данных»</p> <p>Тема 1.1 Введение в статистические методы анализа данных.</p> <p>Тема 1.2 Методы первичной обработки и представления данных.</p> <p>Тема 1.3 Методы выявления и устранения аномальных наблюдений.</p> <p>Раздел №2 «Множественный корреляционно-регрессионный анализ. Классическая модель множественной регрессии»</p> <p>Тема 2.1 Корреляционно-регрессионный анализ. Модель множественной регрессии.</p> <p>Тема 2.2 Фиктивные переменные. Критерий Чоу проверки однородности.</p> <p>Тема 2.3 Множественные коэффициенты корреляции.</p> <p>Тема 2.4 Статистический анализ адекватности модели множественной регрессии.</p> <p>Тема 2.5 Мультиколлинеарность факторов.</p> <p>Раздел №3 «Обобщенная линейная модель множественной регрессии. Регрессионные модели с автокоррелированными остатками»</p> <p>Тема 3.1 Обобщенная линейная</p>	<ol style="list-style-type: none"> 1. Основные этапы эмпирического исследования. 2. Дискретные, непрерывные данные. Пространственная выборка, временной ряд, панельные данные. 3. Аномальные наблюдения. Критерии проверки аномальности наблюдений. 4. Общий вид модели множественной регрессии. Линейная модель множественной регрессии. Матричная форма. 5. Основные предпосылки регрессионного анализа. Метод наименьших квадратов. МНК-оценки параметров множественной линейной регрессии. Теорема Гаусса-Маркова. 6. Статистические свойства МНК оценок. Проверка гипотезы о статистической значимости теоретических коэффициентов регрессии. Доверительные интервалы для теоретических коэффициентов. 7. Частный коэффициент корреляции, множественный коэффициент корреляции. Проверка гипотезы о статистической значимости частного коэффициента корреляции. 8. Коэффициент детермина- 	<p>Устный контроль/ опрос на семинарских занятиях, зачете, экзамене; анализ докладов на семинарских занятиях; анализ защиты рефератов; анализ защиты проектов; применение теоретических знаний при анализе (разборе) конкретных практико-ориентированных ситуаций и профессионально-прикладных задач, анализ использования теоретических знаний в процессе решения кейсов, в ходе деловых игр; письменный контроль, анализ содержания эссе; тестирование (выполнение тестовых заданий)</p>

	<p>модель множественной регрессии. Обобщенный МНК Тема 3.2 Гетероскедастичность. Методы выявления гетероскедастичности. Тема 3.3 Взвешенный МНК. Устранение гетероскедастичности. Тема 3.4 Регрессионные модели с автокоррелированными остатками. Тема 3.5 Устранение автокорреляции Раздел №4 «Кластерный анализ» Тема 4.1 Кластерный анализ. Основные этапы кластерного анализа. Тема 4.2 Метрики расстояний. Свойства кластеров. Методы кластерного анализа. Тема 4.3 Расстояния между кластерами. Методы группировки кластеров. Тема 4.4 Иерархические методы кластеризации. Агломеративные методы. Тема 4.5 Иерархические дивизимные методы кластеризации. Тема 4.6 Примеры реализации иерархических методов кластеризации. Тема 4.7 Итеративные методы. Метод k-средних. Тема 4.8 Метод k-средних. Критерии качества кластеризации. Тема 4.9 Метод нечетких k-средних Раздел №5 «Методы классификации» Тема 5.1 Понятие методов классификации. Тема 5.2 Логистическая регрессия. Тема 5.3 Метрики качества классификации. Тема 5.4 Метрики качества классификации. Регуляризация в задаче регрессии. Тема 5.5 Методы деревьев решений. Тема 5.6 Ансамбли деревьев решений: Случайный лес. Тема 5.7 Ансамбли деревьев решений: Градиентный бустинг.</p>	<p>ции. Свойства. Дисперсионное отношение Фишера. Проверка гипотезы о статистической значимости уравнения регрессии. 9. Обобщенная линейная модель множественной регрессии. Обобщенный МНК. 10. Гомоскедастичность, гетероскедастичность. Критерии проверки остатков на гетероскедастичность. Методы устранения гетероскедастичности. 11. Автокорреляция остатков. Критерии проверки автокорреляции. Устранение автокорреляции. 12. Понятие кластерного анализа. Основные этапы кластерного анализа. 13. Основные меры сходства объектов. Расстояние между объектами: основные виды, формулы расчета. 14. Свойства кластера: плотность, дисперсия, размеры, форма, делимость. Расстояние между кластерами: основные виды, формулы расчета. 15. Иерархические агломеративные методы: понятие, алгоритм реализации. 16. Иерархические дивизимные методы: понятие, алгоритм реализации. 17. Метод k-средних: понятие, алгоритм реализации. Функционалы качества разбиения объектов методом k-средних. 18. Метод нечетких k-средних: понятие, алгоритм реализации. Функционалы качества разбиения объектов методом k-средних. 19. Постановка задачи классификации. Основные методы классификации. 20. Общий вид логистической регрессии. Оценка логистической регрессии. Понятие сбалансированной выборки. 21. Метрики качества классификации. 22. Методы деревьев решений. 23. Понятие ансамблей решений. Случайный лес. Градиентный бустинг.</p>	
<p>ИПК-2.2. Умеет анализировать входные данные; планировать работы.</p>	<p>Раздел 1 «Введение в статистические методы анализа данных» Тема 1.1 Введение в статистические методы анализа данных. Тема 1.2 Методы первичной обработки и представления данных. Тема 1.3 Методы выявления и устранения аномальных наблюдений. Раздел №2 «Множественный</p>	<p>1. Основные этапы эмпирического исследования. 2. Дискретные, непрерывные данные. Пространственная выборка, временной ряд, панельные данные. 3. Аномальные наблюдения. Критерии проверки аномальности наблюдений. 4. Общий вид модели множественной регрессии. Линей-</p>	<p>Анализ проявленных умений при решении кейсов, в ходе деловых игр; письменный контроль, анализ качества решений профессиональных задач в контрольных работах; ана-</p>

	<p>корреляционно-регрессионный анализ. Классическая модель множественной регрессии» Тема 2.1 Корреляционно-регрессионный анализ. Модель множественной регрессии. Тема 2.2 Фиктивные переменные. Критерий Чоу проверки однородности. Тема 2.3 Множественные коэффициенты корреляции. Тема 2.4 Статистический анализ адекватности модели множественной регрессии. Тема 2.5 Мультиколлинеарность факторов. Раздел №3 «Обобщенная линейная модель множественной регрессии. Регрессионные модели с автокоррелированными остатками» Тема 3.1 Обобщенная линейная модель множественной регрессии. Обобщенный МНК Тема 3.2 Гетероскедастичность. Методы выявления гетероскедастичности. Тема 3.3 Взвешенный МНК. Устранение гетероскедастичности. Тема 3.4 Регрессионные модели с автокоррелированными остатками. Тема 3.5 Устранение автокорреляции Раздел №4 «Кластерный анализ» Тема 4.1 Кластерный анализ. Основные этапы кластерного анализа. Тема 4.2 Метрики расстояний. Свойства кластеров. Методы кластерного анализа. Тема 4.3 Расстояния между кластерами. Методы группировки кластеров. Тема 4.4 Иерархические методы кластеризации. Агломеративные методы. Тема 4.5 Иерархические дивизимные методы кластеризации. Тема 4.6 Примеры реализации иерархических методов кластеризации. Тема 4.7 Итеративные методы. Метод k-средних. Тема 4.8 Метод k-средних. Критерии качества кластеризации. Тема 4.9 Метод нечетких k-средних Раздел №5 «Методы классификации» Тема 5.1 Понятие методов классификации. Тема 5.2 Логистическая регрессия.</p>	<p>ная модель множественной регрессии. Матричная форма. 5. Основные предпосылки регрессионного анализа. Метод наименьших квадратов. МНК-оценки параметров множественной линейной регрессии. Теорема Гаусса-Маркова. 6. Статистические свойства МНК оценок. Проверка гипотезы о статистической значимости теоретических коэффициентов регрессии. Доверительные интервалы для теоретических коэффициентов. 7. Частный коэффициент корреляции, множественный коэффициент корреляции. Проверка гипотезы о статистической значимости частного коэффициента корреляции. 8. Коэффициент детерминации. Свойства. Дисперсионное отношение Фишера. Проверка гипотезы о статистической значимости уравнения регрессии. 9. Обобщенная линейная модель множественной регрессии. Обобщенный МНК. 10. Гомоскедастичность, гетероскедастичность. Критерии проверки остатков на гетероскедастичность. Методы устранения гетероскедастичности. 11. Автокорреляция остатков. Критерии проверки автокорреляции. Устранение автокорреляции. 12. Понятие кластерного анализа. Основные этапы кластерного анализа. 13. Основные меры сходства объектов. Расстояние между объектами: основные виды, формулы расчета. 14. Свойства кластера: плотность, дисперсия, размеры, форма, делимость. Расстояние между кластерами: основные виды, формулы расчета. 15. Иерархические агломеративные методы: понятие, алгоритм реализации. 16. Иерархические дивизимные методы: понятие, алгоритм реализации. 17. Метод k-средних: понятие, алгоритм реализации. Функционалы качества разбиения объектов методом k-средних. 18. Метод нечетких k-средних: понятие, алгоритм реализации. Функционалы качества разбиения объектов методом k-средних. 19. Постановка задачи классификации. Основные методы</p>	<p>лиз содержания профессионально-ориентированных эссе; тестирование (выполнение тестовых заданий); анализ защит профессионально-ориентированных проектов; опрос на семинарских занятиях, зачете, анализ докладов на семинарских занятиях; анализ защиты рефератов; анализ решения конкретных практико-ориентированных ситуаций и профессионально-прикладных задач, анализ выполнения контрольных работ</p>
--	--	--	---

	<p>Тема 5.3 Метрики качества классификации.</p> <p>Тема 5.4 Метрики качества классификации. Регуляризация в задаче регрессии.</p> <p>Тема 5.5 Методы деревьев решений.</p> <p>Тема 5.6 Ансамбли деревьев решений: Случайный лес.</p> <p>Тема 5.7 Ансамбли деревьев решений: Градиентный бустинг.</p>	<p>классификации.</p> <p>20. Общий вид логистической регрессии. Оценка логистической регрессии. Понятие сбалансированной выборки.</p> <p>21. Метрики качества классификации.</p> <p>22. Методы деревьев решений.</p> <p>23. Понятие ансамблей решений. Случайный лес. Градиентный бустинг.</p>	
<p>ИПК-2.3. Владеть системным подходом для выбора технологии управления требованиями; представлениями исходных данных для разработки плана управления требованиями; системой согласования в части инженерно-технологического обеспечения плана управления требованиями с заинтересованными сторонами</p>	<p>Раздел 1 «Введение в статистические методы анализа данных»</p> <p>Тема 1.1 Введение в статистические методы анализа данных.</p> <p>Тема 1.2 Методы первичной обработки и представления данных.</p> <p>Тема 1.3 Методы выявления и устранения аномальных наблюдений.</p> <p>Раздел №2 «Множественный корреляционно-регрессионный анализ. Классическая модель множественной регрессии»</p> <p>Тема 2.1 Корреляционно-регрессионный анализ. Модель множественной регрессии.</p> <p>Тема 2.2 Фиктивные переменные. Критерий Чоу проверки однородности.</p> <p>Тема 2.3 Множественные коэффициенты корреляции.</p> <p>Тема 2.4 Статистический анализ адекватности модели множественной регрессии.</p> <p>Тема 2.5 Мультиколлинеарность факторов.</p> <p>Раздел №3 «Обобщенная линейная модель множественной регрессии. Регрессионные модели с автокоррелированными остатками»</p> <p>Тема 3.1 Обобщенная линейная модель множественной регрессии. Обобщенный МНК</p> <p>Тема 3.2 Гетероскедастичность. Методы выявления гетероскедастичности.</p> <p>Тема 3.3 Взвешенный МНК. Устранение гетероскедастичности.</p> <p>Тема 3.4 Регрессионные модели с автокоррелированными остатками.</p> <p>Тема 3.5 Устранение автокорреляции</p> <p>Раздел №4 «Кластерный анализ»</p> <p>Тема 4.1 Кластерный анализ. Основные этапы кластерного анализа.</p> <p>Тема 4.2 Метрики расстояний. Свойства кластеров. Методы кластерного анализа.</p> <p>Тема 4.3 Расстояния между кластерами. Методы группировки</p>	<ol style="list-style-type: none"> 1. Основные этапы эмпирического исследования. 2. Дискретные, непрерывные данные. Пространственная выборка, временной ряд, панельные данные. 3. Аномальные наблюдения. Критерии проверки аномальности наблюдений. 4. Общий вид модели множественной регрессии. Линейная модель множественной регрессии. Матричная форма. 5. Основные предпосылки регрессионного анализа. Метод наименьших квадратов. МНК-оценки параметров множественной линейной регрессии. Теорема Гаусса-Маркова. 6. Статистические свойства МНК оценок. Проверка гипотезы о статистической значимости теоретических коэффициентов регрессии. Доверительные интервалы для теоретических коэффициентов. 7. Частный коэффициент корреляции, множественный коэффициент корреляции. Проверка гипотезы о статистической значимости частного коэффициента корреляции. 8. Коэффициент детерминации. Свойства. Дисперсионное отношение Фишера. Проверка гипотезы о статистической значимости уравнения регрессии. 9. Обобщенная линейная модель множественной регрессии. Обобщенный МНК. 10. Гомоскедастичность, гетероскедастичность. Критерии проверки остатков на гетероскедастичность. Методы устранения гетероскедастичности. 11. Автокорреляция остатков. Критерии проверки автокорреляции. Устранение автокорреляции. 12. Понятие кластерного анализа. Основные этапы кластерного анализа. 13. Основные меры сходства объектов. Расстояние между объектами: основные виды, формулы расчета. 	<p>Анализ проявленных навыков при решении кейсов, в ходе деловых игр; письменный контроль, анализ качества решений профессиональных задач в контрольных работах; анализ содержания профессионально-ориентированных эссе; тестирование (выполнение тестовых заданий); анализ защит профессионально-ориентированных проектов; опрос на семинарских занятиях, зачете, экзамене; анализ докладов на семинарских занятиях; анализ защиты рефератов; анализ решения конкретных практико-ориентированных ситуаций и профессионально-прикладных задач, анализ выполнения контрольных работ</p>

	<p>кластеров. Тема 4.4 Иерархические методы кластеризации. Агломеративные методы. Тема 4.5 Иерархические дивизимные методы кластеризации. Тема 4.6 Примеры реализации иерархических методов кластеризации. Тема 4.7 Итеративные методы. Метод к-средних. Тема 4.8 Метод к-средних. Критерии качества кластеризации. Тема 4.9 Метод нечетких к-средних Раздел №5 «Методы классификации» Тема 5.1 Понятие методов классификации. Тема 5.2 Логистическая регрессия. Тема 5.3 Метрики качества классификации. Тема 5.4 Метрики качества классификации. Регуляризация в задаче регрессии. Тема 5.5 Методы деревьев решений. Тема 5.6 Ансамбли деревьев решений: Случайный лес. Тема 5.7 Ансамбли деревьев решений: Градиентный бустинг.</p>	<p>14. Свойства кластера: плотность, дисперсия, размеры, форма, отделимость. Расстояние между кластерами: основные виды, формулы расчета. 15. Иерархические агломеративные методы: понятие, алгоритм реализации. 16. Иерархические дивизимные методы: понятие, алгоритм реализации. 17. Метод к-средних: понятие, алгоритм реализации. Функционалы качества разбиения объектов методом к-средних. 18. Метод нечетких к-средних: понятие, алгоритм реализации. Функционалы качества разбиения объектов методом к-средних. 19. Постановка задачи классификации. Основные методы классификации. 20. Общий вид логистической регрессии. Оценка логистической регрессии. Понятие сбалансированной выборки. 21. Метрики качества классификации. 22. Методы деревьев решений. 23. Понятие ансамблей решений. Случайный лес. Градиентный бустинг.</p>	
<p>ПК-3. Способность применять системный подход для составления технико-экономического обоснования проектных решений и технического задания на разработку информационной системы</p>			
<p>ИПКЗ.1. Знает возможности типовой ИС; предметную область автоматизации; методы выявления требований; технологии межличностной и групповой коммуникации в деловом взаимодействии, основы конфликтологии; технологии подготовки и проведения презентаций; архитектуру, устройство и функционирование вычислительных систем; коммуникационное оборудование; сетевые протоколы; основы современных операционных систем; основы современных систем управления базами данных; устройство и функционирование современных ИС; современные стандарты информационного взаимодействия систем; программные средства и платформы инфраструктуры информационных технологий организаций; современные подходы и стандарты автоматизации организации (например, CRM, MRP, ERP..., ITIL, ITSM); основы теории систем</p>	<p>Раздел 1 «Введение в статистические методы анализа данных» Тема 1.1 Введение в статистические методы анализа данных. Тема 1.2 Методы первичной обработки и представления данных. Тема 1.3 Методы выявления и устранения аномальных наблюдений. Раздел №2 «Множественный корреляционно-регрессионный анализ. Классическая модель множественной регрессии» Тема 2.1 Корреляционно-регрессионный анализ. Модель множественной регрессии. Тема 2.2 Фиктивные переменные. Критерий Чоу проверки однородности. Тема 2.3 Множественные коэффициенты корреляции. Тема 2.4 Статистический анализ адекватности модели множественной регрессии. Тема 2.5 Мультиколлинеарность факторов. Раздел №3 «Обобщенная линейная модель множественной регрессии. Регрессионные модели с автокоррелированными остатками»</p>	<p>1. Основные этапы эмпирического исследования. 2. Дискретные, непрерывные данные. Пространственная выборка, временной ряд, панельные данные. 3. Аномальные наблюдения. Критерии проверки аномальности наблюдений. 4. Общий вид модели множественной регрессии. Линейная модель множественной регрессии. Матричная форма. 5. Основные предпосылки регрессионного анализа. Метод наименьших квадратов. МНК-оценки параметров множественной линейной регрессии. Теорема Гаусса-Маркова. 6. Статистические свойства МНК оценок. Проверка гипотезы о статистической значимости теоретических коэффициентов регрессии. Доверительные интервалы для теоретических коэффициентов. 7. Частный коэффициент корреляции, множественный коэффициент корреляции. Проверка гипотезы о статистической значимости частного коэффициента корреляции.</p>	<p>Устный контроль/ опрос на семинарских занятиях, зачете, экзамене; анализ докладов на семинарских занятиях; анализ защиты рефератов; анализ защиты проектов; применение теоретических знаний при анализе (разборе) конкретных практико-ориентированных ситуаций и профессионально-прикладных задач, анализ использования теоретических знаний в процессе решения кейсов, в ходе деловых игр; письменный контроль, анализ содержания эссе; тестирование (выполнение тестовых заданий)</p>

<p>и системного анализа; методики описания и моделирования бизнес-процессов, средства моделирования бизнес-процессов; системы классификации и кодирования информации, в том числе присвоение кодов документам и элементам справочников; отраслевая нормативная техническая документация; источники информации, необходимой для профессиональной деятельности; современный отечественный и зарубежный опыт в профессиональной деятельности; формирование и механизмы рыночных процессов организации; основы менеджмента качества; основы бухгалтерского учета и отчетности организаций; основы налогового законодательства Российской Федерации; основы управленческого учета; основы финансового учета и бюджетирования; основы международных стандартов финансовой отчетности (МСФО); основы управления торговлей, поставками и запасами; основы организации производства; основы управления взаимоотношениями с клиентами и заказчиками (CRM); основы управления персоналом, включая вопросы оплаты труда; основы теории управления; современные инструменты и методы управления организацией, в том числе методы планирования деятельности, распределения поручений, контроля исполнения, принятия решений; методология ведения документооборота в организациях; инструменты и методы определения финансовых и производственных показателей деятельности организаций; основы организационной диагностики; инструменты и методы моделирования бизнес-процессов организации; основы реинжиниринга бизнес-процессов организации; управление содержанием проекта документирование требований, анализ продукта, модерируемые совещания; культура речи; правила деловой переписки.</p>	<p>Тема 3.1 Обобщенная линейная модель множественной регрессии. Обобщенный МНК Тема 3.2 Гетероскедастичность. Методы выявления гетероскедастичности. Тема 3.3 Взвешенный МНК. Устранение гетероскедастичности. Тема 3.4 Регрессионные модели с автокоррелированными остатками. Тема 3.5 Устранение автокорреляции Раздел №4 «Кластерный анализ» Тема 4.1 Кластерный анализ. Основные этапы кластерного анализа. Тема 4.2 Метрики расстояний. Свойства кластеров. Методы кластерного анализа. Тема 4.3 Расстояния между кластерами. Методы группировки кластеров. Тема 4.4 Иерархические методы кластеризации. Агломеративные методы. Тема 4.5 Иерархические дивизимные методы кластеризации. Тема 4.6 Примеры реализации иерархических методов кластеризации. Тема 4.7 Итеративные методы. Метод k-средних. Тема 4.8 Метод k-средних. Критерии качества кластеризации. Тема 4.9 Метод нечетких k-средних Раздел №5 «Методы классификации» Тема 5.1 Понятие методов классификации. Тема 5.2 Логистическая регрессия. Тема 5.3 Метрики качества классификации. Тема 5.4 Метрики качества классификации. Регуляризация в задаче регрессии. Тема 5.5 Методы деревьев решений. Тема 5.6 Ансамбли деревьев решений: Случайный лес. Тема 5.7 Ансамбли деревьев решений: Градиентный бустинг.</p>	<p>8. Коэффициент детерминации. Свойства. Дисперсионное отношение Фишера. Проверка гипотезы о статистической значимости уравнения регрессии. 9. Обобщенная линейная модель множественной регрессии. Обобщенный МНК. 10. Гомоскедастичность, гетероскедастичность. Критерии проверки остатков на гетероскедастичность. Методы устранения гетероскедастичности. 11. Автокорреляция остатков. Критерии проверки автокорреляции. Устранение автокорреляции. 12. Понятие кластерного анализа. Основные этапы кластерного анализа. 13. Основные меры сходства объектов. Расстояние между объектами: основные виды, формулы расчета. 14. Свойства кластера: плотность, дисперсия, размеры, форма, делимость. Расстояние между кластерами: основные виды, формулы расчета. 15. Иерархические агломеративные методы: понятие, алгоритм реализации. 16. Иерархические дивизимные методы: понятие, алгоритм реализации. 17. Метод k-средних: понятие, алгоритм реализации. Функционалы качества разбиения объектов методом k-средних. 18. Метод нечетких k-средних: понятие, алгоритм реализации. Функционалы качества разбиения объектов методом k-средних. 19. Постановка задачи классификации. Основные методы классификации. 20. Общий вид логистической регрессии. Оценка логистической регрессии. Понятие сбалансированной выборки. 21. Метрики качества классификации. 22. Методы деревьев решений. 23. Понятие ансамблей решений. Случайный лес. Градиентный бустинг.</p>	
---	---	--	--

<p>ИПК-3.2 Умеет проводить переговоры; проводить презентации; подготавливать протоколы мероприятий.</p>	<p>Раздел 1 «Введение в статистические методы анализа данных» Тема 1.1 Введение в статистические методы анализа данных. Тема 1.2 Методы первичной обработки и представления данных. Тема 1.3 Методы выявления и устранения аномальных наблюдений. Раздел №2 «Множественный корреляционно-регрессионный анализ. Классическая модель множественной регрессии» Тема 2.1 Корреляционно-регрессионный анализ. Модель множественной регрессии. Тема 2.2 Фиктивные переменные. Критерий Чоу проверки однородности. Тема 2.3 Множественные коэффициенты корреляции. Тема 2.4 Статистический анализ адекватности модели множественной регрессии. Тема 2.5 Мультиколлинеарность факторов. Раздел №3 «Обобщенная линейная модель множественной регрессии. Регрессионные модели с автокоррелированными остатками» Тема 3.1 Обобщенная линейная модель множественной регрессии. Обобщенный МНК Тема 3.2 Гетероскедастичность. Методы выявления гетероскедастичности. Тема 3.3 Взвешенный МНК. Устранение гетероскедастичности. Тема 3.4 Регрессионные модели с автокоррелированными остатками. Тема 3.5 Устранение автокорреляции Раздел №4 «Кластерный анализ» Тема 4.1 Кластерный анализ. Основные этапы кластерного анализа. Тема 4.2 Метрики расстояний. Свойства кластеров. Методы кластерного анализа. Тема 4.3 Расстояния между кластерами. Методы группировки кластеров. Тема 4.4 Иерархические методы кластеризации. Агломеративные методы. Тема 4.5 Иерархические дивизимные методы кластеризации. Тема 4.6 Примеры реализации иерархических методов кластеризации. Тема 4.7 Итеративные методы. Метод k-средних.</p>	<ol style="list-style-type: none"> 1. Основные этапы эмпирического исследования. 2. Дискретные, непрерывные данные. Пространственная выборка, временной ряд, панельные данные. 3. Аномальные наблюдения. Критерии проверки аномальности наблюдений. 4. Общий вид модели множественной регрессии. Линейная модель множественной регрессии. Матричная форма. 5. Основные предпосылки регрессионного анализа. Метод наименьших квадратов. МНК-оценки параметров множественной линейной регрессии. Теорема Гаусса-Маркова. 6. Статистические свойства МНК оценок. Проверка гипотезы о статистической значимости теоретических коэффициентов регрессии. Доверительные интервалы для теоретических коэффициентов. 7. Частный коэффициент корреляции, множественный коэффициент корреляции. Проверка гипотезы о статистической значимости частного коэффициента корреляции. 8. Коэффициент детерминации. Свойства. Дисперсионное отношение Фишера. Проверка гипотезы о статистической значимости уравнения регрессии. 9. Обобщенная линейная модель множественной регрессии. Обобщенный МНК. 10. Гомоскедастичность, гетероскедастичность. Критерии проверки остатков на гетероскедастичность. Методы устранения гетероскедастичности. 11. Автокорреляция остатков. Критерии проверки автокорреляции. Устранение автокорреляции. 12. Понятие кластерного анализа. Основные этапы кластерного анализа. 13. Основные меры сходства объектов. Расстояние между объектами: основные виды, формулы расчета. 14. Свойства кластера: плотность, дисперсия, размеры, форма, отделимость. Расстояние между кластерами: основные виды, формулы расчета. 15. Иерархические агломеративные методы: понятие, алгоритм реализации. 16. Иерархические дивизимные методы: понятие, алгоритм реализации. 17. Метод k-средних: поня- 	<p>Анализ проявленных умений при решении кейсов, в ходе деловых игр; письменный контроль, анализ качества решений профессиональных задач в контрольных работах; анализ содержания профессионально-ориентированных эссе; тестирование (выполнение тестовых заданий); анализ защит профессионально-ориентированных проектов; опрос на семинарских занятиях, зачете, анализ докладов на семинарских занятиях; анализ защиты рефератов; анализ решения конкретных практико-ориентированных ситуаций и профессионально-прикладных задач, анализ выполнения контрольных работ</p>
---	---	--	---

	<p>Тема 4.8 Метод к-средних. Критерии качества кластеризации.</p> <p>Тема 4.9 Метод нечетких к-средних</p> <p>Раздел №5 «Методы классификации»</p> <p>Тема 5.1 Понятие методов классификации.</p> <p>Тема 5.2 Логистическая регрессия.</p> <p>Тема 5.3 Метрики качества классификации.</p> <p>Тема 5.4 Метрики качества классификации. Регуляризация в задаче регрессии.</p> <p>Тема 5.5 Методы деревьев решений.</p> <p>Тема 5.6 Ансамбли деревьев решений: Случайный лес.</p> <p>Тема 5.7 Ансамбли деревьев решений: Градиентный бустинг.</p>	<p>тие, алгоритм реализации. Функционалы качества разбиения объектов методом к-средних.</p> <p>18. Метод нечетких к-средних: понятие, алгоритм реализации. Функционалы качества разбиения объектов методом к-средних.</p> <p>19. Постановка задачи классификации. Основные методы классификации.</p> <p>20. Общий вид логистической регрессии. Оценка логистической регрессии. Понятие сбалансированной выборки.</p> <p>21. Метрики качества классификации.</p> <p>22. Методы деревьев решений.</p> <p>23. Понятие ансамблей решений. Случайный лес. Градиентный бустинг.</p>	
<p>ИПК 3.3. Владеет основами выявления первоначальных требований заказчика к ИС; системой информирования заказчика о возможностях типовой ИС и вариантах ее модификации; определениями возможностей достижения соответствия ИС первоначальным требованиям заказчика; основами составления протоколов переговоров с заказчиком</p>	<p>Раздел 1 «Введение в статистические методы анализа данных»</p> <p>Тема 1.1 Введение в статистические методы анализа данных.</p> <p>Тема 1.2 Методы первичной обработки и представления данных.</p> <p>Тема 1.3 Методы выявления и устранения аномальных наблюдений.</p> <p>Раздел №2 «Множественный корреляционно-регрессионный анализ. Классическая модель множественной регрессии»</p> <p>Тема 2.1 Корреляционно-регрессионный анализ. Модель множественной регрессии.</p> <p>Тема 2.2 Фиктивные переменные. Критерий Чоу проверки однородности.</p> <p>Тема 2.3 Множественные коэффициенты корреляции.</p> <p>Тема 2.4 Статистический анализ адекватности модели множественной регрессии.</p> <p>Тема 2.5 Мультиколлинеарность факторов.</p> <p>Раздел №3 «Обобщенная линейная модель множественной регрессии. Регрессионные модели с автокоррелированными остатками»</p> <p>Тема 3.1 Обобщенная линейная модель множественной регрессии. Обобщенный МНК</p> <p>Тема 3.2 Гетероскедастичность. Методы выявления гетероскедастичности.</p> <p>Тема 3.3 Взвешенный МНК. Устранение гетероскедастичности.</p> <p>Тема 3.4 Регрессионные модели с автокоррелированными остатками.</p> <p>Тема 3.5 Устранение автокорре-</p>	<ol style="list-style-type: none"> 1. Основные этапы эмпирического исследования. 2. Дискретные, непрерывные данные. Пространственная выборка, временной ряд, панельные данные. 3. Аномальные наблюдения. Критерии проверки аномальности наблюдений. 4. Общий вид модели множественной регрессии. Линейная модель множественной регрессии. Матричная форма. 5. Основные предпосылки регрессионного анализа. Метод наименьших квадратов. МНК-оценки параметров множественной линейной регрессии. Теорема Гаусса-Маркова. 6. Статистические свойства МНК оценок. Проверка гипотезы о статистической значимости теоретических коэффициентов регрессии. Доверительные интервалы для теоретических коэффициентов. 7. Частный коэффициент корреляции, множественный коэффициент корреляции. Проверка гипотезы о статистической значимости частного коэффициента корреляции. 8. Коэффициент детерминации. Свойства. Дисперсионное отношение Фишера. Проверка гипотезы о статистической значимости уравнения регрессии. 9. Обобщенная линейная модель множественной регрессии. Обобщенный МНК. 10. Гомоскедастичность, гетероскедастичность. Критерии проверки остатков на гетероскедастичность. Методы устранения гетероскедастичности. 	<p>Анализ проявленных навыков при решении кейсов, в ходе деловых игр; письменный контроль, анализ качества решений профессиональных задач в контрольных работах; анализ содержания профессионально-ориентированных эссе; тестирование (выполнение тестовых заданий); анализ защит профессионально-ориентированных проектов; опрос на семинарских занятиях, зачете, экзамене; анализ докладов на семинарских занятиях; анализ защиты рефератов; анализ решения конкретных практико-ориентированных ситуаций и профессионально-прикладных задач, анализ выполнения контрольных работ</p>

	<p>ляции</p> <p>Раздел №4 «Кластерный анализ»</p> <p>Тема 4.1 Кластерный анализ. Основные этапы кластерного анализа.</p> <p>Тема 4.2 Метрики расстояний. Свойства кластеров. Методы кластерного анализа.</p> <p>Тема 4.3 Расстояния между кластерами. Методы группировки кластеров.</p> <p>Тема 4.4 Иерархические методы кластеризации. Агломеративные методы.</p> <p>Тема 4.5 Иерархические дивизимные методы кластеризации.</p> <p>Тема 4.6 Примеры реализации иерархических методов кластеризации.</p> <p>Тема 4.7 Итеративные методы. Метод k-средних.</p> <p>Тема 4.8 Метод k-средних. Критерии качества кластеризации.</p> <p>Тема 4.9 Метод нечетких k-средних</p> <p>Раздел №5 «Методы классификации»</p> <p>Тема 5.1 Понятие методов классификации.</p> <p>Тема 5.2 Логистическая регрессия.</p> <p>Тема 5.3 Метрики качества классификации.</p> <p>Тема 5.4 Метрики качества классификации. Регуляризация в задаче регрессии.</p> <p>Тема 5.5 Методы деревьев решений.</p> <p>Тема 5.6 Ансамбли деревьев решений: Случайный лес.</p> <p>Тема 5.7 Ансамбли деревьев решений: Градиентный бустинг.</p>	<p>11. Автокорреляция остатков. Критерии проверки автокорреляции. Устранение автокорреляции.</p> <p>12. Понятие кластерного анализа. Основные этапы кластерного анализа.</p> <p>13. Основные меры сходства объектов. Расстояние между объектами: основные виды, формулы расчета.</p> <p>14. Свойства кластера: плотность, дисперсия, размеры, форма, отделимость. Расстояние между кластерами: основные виды, формулы расчета.</p> <p>15. Иерархические агломеративные методы: понятие, алгоритм реализации.</p> <p>16. Иерархические дивизимные методы: понятие, алгоритм реализации.</p> <p>17. Метод k-средних: понятие, алгоритм реализации. Функционалы качества разбиения объектов методом k-средних.</p> <p>18. Метод нечетких k-средних: понятие, алгоритм реализации. Функционалы качества разбиения объектов методом k-средних.</p> <p>19. Постановка задачи классификации. Основные методы классификации.</p> <p>20. Общий вид логистической регрессии. Оценка логистической регрессии. Понятие сбалансированной выборки.</p> <p>21. Метрики качества классификации.</p> <p>22. Методы деревьев решений.</p> <p>23. Понятие ансамблей решений. Случайный лес. Градиентный бустинг.</p>	
--	--	---	--

7.2.Перечень вопросов для подготовки к промежуточной аттестации (экзамен)

1. Основные этапы эмпирического исследования.
2. Дискретные, непрерывные данные. Пространственная выборка, временной ряд, панельные данные.
3. Аномальные наблюдения. Критерии проверки аномальности наблюдений.
4. Общий вид модели множественной регрессии. Линейная модель множественной регрессии. Матричная форма.
5. Основные предпосылки регрессионного анализа. Метод наименьших квадратов. МНК-оценки параметров множественной линейной регрессии. Теорема Гаусса-Маркова.
6. Статистические свойства МНК оценок. Проверка гипотезы о статистической значимости теоретических коэффициентов регрессии. Доверительные интервалы для теоретических коэффициентов.
7. Частный коэффициент корреляции, множественный коэффициент корреляции. Проверка гипотезы о статистической значимости частного коэффициента корреляции.
8. Коэффициент детерминации. Свойства. Дисперсионное отношение Фишера. Проверка гипотезы о статистической значимости уравнения регрессии.
9. Обобщенная линейная модель множественной регрессии. Обобщенный МНК.
10. Гомоскедастичность, гетероскедастичность. Критерии проверки остатков на гетероскедастичность. Методы устранения гетероскедастичности.

11. Автокорреляция остатков. Критерии проверки автокорреляции. Устранение автокорреляции.
12. Понятие кластерного анализа. Основные этапы кластерного анализа.
13. Основные меры сходства объектов. Расстояние между объектами: основные виды, формулы расчета.
14. Свойства кластера: плотность, дисперсия, размеры, форма, отделимость. Расстояние между кластерами: основные виды, формулы расчета.
15. Иерархические агломеративные методы: понятие, алгоритм реализации.
16. Иерархические дивизимные методы: понятие, алгоритм реализации.
17. Метод k -средних: понятие, алгоритм реализации. Функционалы качества разбиения объектов методом k -средних.
18. Метод нечетких k -средних: понятие, алгоритм реализации. Функционалы качества разбиения объектов методом k -средних.
19. Постановка задачи классификации. Основные методы классификации.
20. Общий вид логистической регрессии. Оценка логистической регрессии. Понятие сбалансированной выборки.
21. Метрики качества классификации.
22. Методы деревьев решений.
23. Понятие ансамблей решений. Случайный лес. Градиентный бустинг.

7.3. Описание показателей и критериев оценивания сформированности компетенций на различных этапах их формирования; шкалы и процедуры оценивания

7.3.1. Оценивание ответов на вопросы и выполнения заданий для текущей и промежуточной аттестации

При оценке знаний учитывается уровень сформированности компетенций:

1. Уровень усвоения теоретических положений дисциплины, правильность формулировки основных понятий и закономерностей.
2. Уровень знания фактического материала в объеме программы.
3. Логика, структура и грамотность изложения вопроса.
4. Умение связать теорию с практикой.
5. Умение делать обобщения, выводы.

Шкала оценивания на экзамене и зачете с оценкой

Оценка	Критерии выставления оценки
Отлично	Обучающийся должен: - продемонстрировать глубокое и прочное усвоение знаний программного материала; - исчерпывающе, последовательно, грамотно и логически стройно изложить теоретический материал; - правильно формулировать определения; - продемонстрировать умения самостоятельной работы с литературой; - уметь сделать выводы по излагаемому материалу.
Хорошо	Обучающийся должен: - продемонстрировать достаточно полное знание программного материала; - продемонстрировать знание основных теоретических понятий; достаточно последовательно, грамотно и логически стройно излагать материал; - продемонстрировать умение ориентироваться в литературе; - уметь сделать достаточно обоснованные выводы по излагаемому материалу.
Удовлетворительно	Обучающийся должен: - продемонстрировать общее знание изучаемого материала; - показать общее владение понятийным аппаратом дисциплины; - уметь строить ответ в соответствии со структурой излагаемого вопроса;

	- знать основную рекомендуемую программой учебную литературу.
Неудовлетворительно	Обучающийся демонстрирует: <ul style="list-style-type: none"> - незнание значительной части программного материала; - не владение понятийным аппаратом дисциплины; - существенные ошибки при изложении учебного материала; - неумение строить ответ в соответствии со структурой излагаемого вопроса; - неумение делать выводы по излагаемому материалу.

7.4. МЕТОДИЧЕСКИЕ МАТЕРИАЛЫ, ОПРЕДЕЛЯЮЩИЕ ПРОЦЕДУРЫ ОЦЕНКИ ЗНАНИЙ, УМЕНИЙ, НАВЫКОВ И (ИЛИ) ОПЫТА ДЕЯТЕЛЬНОСТИ, ХАРАКТЕРИЗУЮЩИХ ЭТАПЫ ФОРМИРОВАНИЯ КОМПЕТЕНЦИЙ.

Качество знаний характеризуется способностью обучающегося точно, структурированно и уместно воспроизводить информацию, полученную в процессе освоения дисциплины, в том виде, в котором она была изложена в учебном издании или преподавателем.

Умения, как правило, формируются на занятиях семинарского типа. Задания, направленные на оценку умений, в значительной степени требуют от обучающегося проявления стереотипности мышления, т.е. способности выполнить работу по образцам, с которыми он работал в процессе обучения. Преподаватель же оценивает своевременность и правильность выполнения задания.

Навыки можно трактовать как автоматизированные умения, развитые и закреплённые осознанным самостоятельным трудом. Навыки формируются при самостоятельном выполнении обучающимися практико-ориентированных заданий, моделирующих решение им производственных и социокультурных задач в соответствующей области профессиональной деятельности, как правило, при выполнении домашних заданий, курсовых проектов (работ), научно-исследовательских работ, прохождении практик, при работе индивидуально или в составе группы и т.д.

Устный опрос - это процедура, организованная как специальная беседа преподавателя с группой обучающихся (фронтальный опрос) или с отдельными обучающимися (индивидуальный опрос) с целью оценки сформированности у них основных понятий и усвоения учебного материала. Устный опрос может использоваться как вид контроля и метод оценивания формируемых компетенций (как и качества их формирования) в рамках самых разных форм контроля, таких как: собеседование, коллоквиум, зачет, экзамен по дисциплине. Устный опрос (УО) позволяет оценить знания и кругозор обучающегося, умение логически построить ответ, владение монологической речью и иные коммуникативные навыки. УО обладает большими возможностями воспитательного воздействия преподавателя. Воспитательная функция УО имеет ряд важных аспектов: профессионально-этический и нравственный аспекты, дидактический (систематизация материала при ответе, лучшее запоминание материала при интеллектуальной концентрации), эмоциональный (радость от успешного прохождения собеседования) и др. Обучающая функция УО состоит в выявлении деталей, которые по каким-то причинам оказались недостаточно осмысленными в ходе учебных занятий и при подготовке к зачёту или экзамену. УО обладает также мотивирующей функцией: правильно организованные собеседование, коллоквиум, зачёт и экзамен могут стимулировать учебную деятельность студента, его участие в научной работе.

Тесты являются простейшей формой контроля, направленной на проверку владения терминологическим аппаратом, современными информационными технологиями и конкретными знаниями в области фундаментальных и прикладных дисциплин. Тест может предоставлять возможность выбора из перечня ответов (один или несколько правильных ответов).

Семинарские занятия. Основное назначение семинарских занятий по дисциплине – обеспечить глубокое усвоение обучающимися материалов лекций, прививать навыки самостоятельной работы с литературой, воспитывать умение находить оптимальные решения в условиях изменяющихся отношений, формировать современное профессиональное мышление обучающихся. На семинарских занятиях преподаватель проверяет выполнение самостоятельных заданий и качество усвоения знаний, умений, определяет уровень сформированности компетенций.

Коллоквиум может служить формой не только проверки, но и повышения производительности труда студентов. На коллоквиумах обсуждаются отдельные части, разделы, темы, вопросы изучаемого курса, обычно не включаемые в тематику семинарских и других практических учебных занятий, а также рефераты, проекты и иные работы обучающихся.

Доклад, сообщение - продукт самостоятельной работы студента, представляющий собой публичное выступление по представлению полученных результатов решения определенной учебно-практической, учебно-исследовательской или научной темы.

Контрольная работа - средство проверки умений применять полученные знания для решения задач определенного типа по теме или разделу.

Профессионально-ориентированное эссе – это средство, позволяющее оценить умение обучающегося письменно излагать суть поставленной проблемы, самостоятельно проводить анализ этой проблемы с использованием аналитического инструментария соответствующей дисциплины, делать выводы, обобщающие авторскую позицию по поставленной профессионально-ориентированной проблеме.

Реферат - продукт самостоятельной работы студента, представляющий собой краткое изложение в письменном виде полученных результатов теоретического анализа определенной научной (учебно-исследовательской) темы, где автор раскрывает суть исследуемой проблемы, приводит различные точки зрения, а также собственные взгляды на нее.

Ситуационный анализ - это комплексный анализ ситуации, имевший место в реальной практике профессиональной деятельности специалистов. Комплексный анализ включает в себя следующие составляющие: причинно-следственный анализ (установление причин, которые привели к возникновению данной ситуации, и следствий ее развертывания), системный анализ (определение сущностных предметно-содержательных характеристик, структуры ситуации, ее функций и др.), ценностно-мотивационный анализ (построение системы оценок ситуации, ее составляющих, выявление мотивов, установок, позиций действующих лиц); прогностический анализ (разработка перспектив развития событий по позитивному и негативному сценарию), рекомендательный анализ (выработка рекомендаций относительно поведения действующих лиц ситуации), программно-целевой анализ (разработка программ деятельности для решения данной ситуации).

Творческое задание – это частично регламентированное задание, имеющее нестандартное решение и позволяющее диагностировать умения интегрировать знания различных научных областей, аргументировать собственную точку зрения, доказывать правильность своей позиции. Может выполняться в индивидуальном порядке или группой обучающихся.

Деловая и/или ролевая игра - совместная деятельность группы обучающихся и преподавателя под управлением преподавателя с целью решения учебных и профессионально-ориентированных задач путем игрового моделирования реальной проблемной ситуации. Позволяет оценивать умение анализировать и решать типичные профессиональные задачи.

«Круглый стол», дискуссия – интерактивные оценочные средства, позволяющие включить обучающихся в процесс обсуждения спорного вопроса, проблемы и оценить их умение аргументировать собственную точку зрения. Занятие может проводиться по традиционной (контактной) технологии, либо с использованием телекоммуникационных технологий.

Проект - конечный профессионально-ориентированный продукт, получаемый в результате планирования и выполнения комплекса учебных и исследовательских заданий. Позволяет оценить умения обучающихся самостоятельно конструировать свои знания в процессе решения практических задач и проблем, ориентироваться в информационном пространстве и уровень сформированности аналитических, исследовательских навыков, навыков практического и творческого мышления. Может выполняться в индивидуальном порядке или группой обучающихся.

РАЗДЕЛ 8. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ

При организации процесса изучения дисциплины «Статистические методы анализа данных» обучающемуся рекомендуется придерживаться следующих указаний:

1. Ознакомиться с общим тематическим планом лекционных и лабораторных занятий;
2. Самостоятельно прорабатывать уже пройденный материал лекционных и лабораторных занятий. При необходимости составить список вопросов и обратиться к преподавателю.
3. Перед изучением нового теоретического материала желательно заранее ознакомиться с содержанием предстоящей лекции. При необходимости составить список вопросов и обратиться к преподавателю.
4. Если в ходе рассмотрения нового теоретического или практического материала преподаватель ссылается на полученные ранее знания, умения или навыки, то рекомендуется их повторить.
5. Выполнять индивидуальные задания желательно не только в рамках учебных занятий.
6. Ознакомиться со списком литературы и интернет-ресурсов, рекомендуемых преподавателем для углубленного изучения либо дисциплины в целом, либо отдельных разделов.

Учебно-методическое обеспечение самостоятельной работы студентов по дисциплине включает:

1. Айвазян С.А. Практикум по многомерным статистическим методам: учебное пособие / Айвазян С.А., Мхитарян В.С., Зехин В.А. – Москва: Евразийский открытый институт, Московский государственный университет экономики, статистики и информатики, 2003. – 76 с. – Текст : электронный // IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/10803.html>

2. Кремер Н.Ш. Эконометрика: учебник для студентов вузов / Кремер Н.Ш., Путко Б.А. – Москва: ЮНИТИ-ДАНА, 2017. – 328 с. – ISBN 978-5-238-01720-4. – Текст : электронный // IPR SMART : [сайт]. – URL: <https://www.iprsmarthop.ru/71071.html>

РАЗДЕЛ 9. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Основная литература¹

1. Мхитарян В.С. Эконометрика: учебное пособие / Мхитарян В.С., Архипова М.Ю., Сиротин В.П. – Москва : Евразийский открытый институт, 2012. – 224 с. – ISBN 978-5-374-00053-5. – Текст: электронный // IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/11125.html>

2. Федин, Ф. О. Анализ данных. Часть 1. Подготовка данных к анализу : учебное пособие / Ф. О. Федин, Ф. Ф. Федин. – Москва : Московский городской педагогический университет, 2012. – 204 с. – Текст: электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. — URL: <https://www.iprsmarthop.ru/26444.html>

3. Федин Ф.О. Анализ данных. Часть 2. Инструменты Data Mining: учебное пособие / Федин Ф.О., Федин Ф.Ф. – Москва: Московский городской педагогический университет, 2012. – 308 с. – Текст: электронный // IPR SMART : [сайт]. – URL: <https://www.iprsmarthop.ru/26445.html>

Дополнительная литература²

1. Статистический анализ данных, моделирование и исследование вероятностных закономерностей. Компьютерный подход: монография / Б. Ю. Лемешко, С. Б. Лемешко, С. Н. Постовалов, Е. В. Чимитова. – Новосибирск: Новосибирский государственный технический университет, 2011. – 888 с. – ISBN 978-5-7782-1590-0. – Текст: электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/47719.html>

2. Тимофеев В.С. Эконометрика: учебник / Тимофеев В.С., Фаддеенков А.В., Щеколдин В.Ю.. – Новосибирск: Новосибирский государственный технический университет, 2015. – 352 с. – ISBN 978-5-7782-2658-6. – Текст: электронный // IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/91744.html>

3. Цильковский И.А. Методы анализа знаний и данных: конспект лекций / Цильковский И.А., Волкова В.М.. – Новосибирск: Новосибирский государственный технический университет, 2010. – 68 с. – ISBN 978-57782-1377-7. – Текст: электронный // IPR SMART: [сайт]. – URL: <https://www.iprsmarthop.ru/45385.html>

Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине: интернет-ресурсы, современные профессиональные базы данных, информационные справочные системы

Интернет-ресурсы

URL: <https://www.iprsmarthop.ru/> – электронно-библиотечная система Iprsmart.

Комплект лицензионного и свободно распространяемого программного обеспечения, в том числе отечественного производства

Комплект лицензионного программного обеспечения

Microsoft Open Value Subscription для решений Education Solutions № Tr000544893 от 21.10.2020 г. MDE Windows, Microsoft Office и Office Web Apps. (срок действия до 01.11.2023 г.)

Антивирусное программное обеспечение ESET NOD32 Antivirus Business Edition договор № ИС00-006348 от 14.10.2022 г. (срок действия до 13.10.2025 г.)

Программное обеспечение «Мираполис» система вебинаров - Лицензионный договор 244/09/16-к от 15.09.2016 (Спецификация к Лицензионному договору 244/09/16-к от 15.09.2016, от 11.05.2022 г.) (срок действия до 10.07.2023 г.)

Электронная информационно-образовательная среда «1С: Университет» договор от 10.09.2018 г. №ПРКТ-18281 (бессрочно)

Информационная система «ПервыйБит» сублицензионный оговор от 06.11.2015 г. №009/061115/003 (бессрочно)

¹ Из ЭБС

² Из ЭБС

Система тестирования Indigo лицензионное соглашение (Договор) от 08.11.2018 г. №Д-54792 (бессрочно)

Информационно-поисковая система «Консультант Плюс» - договор об информационно поддержке от 26.12.2014, (бессрочно)

Электронно-библиотечная система IPRsmart лицензионный договор от 01.09.2021 г. №8234/21С (срок действия до 31.08.2024 г.)

Научная электронная библиотека eLIBRARY лицензионный договор SCIENC INDEX № SIO - 3079/2022 от 12.01.2022 г. (срок действия до 27.01.2024 г.)

Свободно распространяемое программное обеспечение

Комплект онлайн сервисов GNU ImageManipulationProgram, свободно распространяемо программное обеспечение

Веб-браузер, Google Ghrome, свободное ПО, ежегодно обновляемое ПО.

Пакет офисных приложений, Office 2016, лицензионное соглашение - Договор №Tr000544893 от 21/10/2020 – 3 года

Пакет офисных приложений, OpenOffice, свободное ПО, ежегодно обновляемое ПО

Просмотр файлов в формате PDF, Adobe Reader, свободно распространяемое ПО, ежегодно обновляемое ПО

Просмотр файлов в формате DJV, WinDjView, свободное ПО, ежегодно обновляемое ПО

Файловый архиватор, 7 Zip, свободное ПО, ежегодно обновляемое ПО

Файловый менеджер, Far, свободно распространяемое ПО, ежегодно обновляемое ПО

Anaconda: дистрибутив языков программирования Python и R.

Программное обеспечение отечественного производства:

Программное обеспечение «Мираполис» система вебинаров - Лицензионный договор 244/09/16-к от 15.09.2016 (Спецификация к Лицензионному договору 244/09/16-к от 15.09.2016, от 11.05.2022 г.) (срок действия до 10.07.2023 г.)

Электронная информационно-образовательная среда «1С: Университет» договор от 10.09.2018 г. №ПРКТ-18281 (бессрочно)

Информационная система «ПервыйБит» сублицензионный договор от 06.11.2015 г. №009/061115/003 (бессрочно)

Система тестирования Indigo лицензионное соглашение (Договор) от 08.11.2018 г. №Д-54792 (бессрочно)

Информационно-поисковая система «Консультант Плюс» - договор МИ-ВИП-79717-56/2022 от 23.12.2021 (срок действия до 31.12.2022 г.)

Информационно-поисковая система «Консультант Плюс» - договор об информационно поддержке от 26.12.2014, (бессрочно)

Электронно-библиотечная система IPRsmart лицензионный договор от 01.09.2021 г. №8234/21С (срок действия до 31.08.2024 г.)

Научная электронная библиотека eLIBRARY лицензионный договор SCIENC INDEX № SIO -3079/2022 от 12.01.2022 г. (срок действия до 27.01.2024 г.)

РАЗДЕЛ 10. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА

Учебная аудитория для проведения занятий лекционного типа, занятий семинарского типа, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации	<u>Оборудование:</u> специализированная мебель (мебель аудиторная (11 столов, 11 стульев, доска аудиторная), стол преподавателя, стул преподавателя). <u>Технические средства обучения:</u> персональный компьютер -11; мультимедийное оборудование (проектор, экран).
Помещение для самостоятельной работы	Специализированная мебель (10 столов, 10 стульев), персональные компьютеры с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду Университета