Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Максимов Алексей Борисович
Должность: директор департамента по образовательной политике
Дата подписания: 31.05.2024 15:12:15
Уникальный программный ключ:
8db180d1a3f02ac9e60521a5672742735c18b1d6

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ
федеральное государственное автономное образовательное учреждение
высшего образования
«МОСКОВСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»
Факультет информационных технологий

УТВЕРЖДАЮ
Декан факультета
«Информационные технологии»
_____ / Д.Г.Демидов /
«15» февраля 2024г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

«Технологии анализа данных»

Направление подготовки/специальность
01.04.02 Прикладная математика и информатика

Профиль/специализация
«Системная аналитика больших данных»

Квалификация
магистр

Формы обучения
очная

Москва, 2024 г.

Разработчик(и):

к.ф-м.н., доцент /А.В. Осипов /

Согласовано:
Заведующий кафедрой «Прикладная информатика»,
к.э.н., доцент /С.В. Суворов /

Содержание

Оглавление

1 Цели, задачи и планируемые результаты обучения по дисциплине 4
2 Место дисциплины в структуре образовательной программы 7
3 Структура и содержание дисциплины 8
3.1 Виды учебной работы и трудоемкость 8
3.2 Тематический план изучения дисциплины 9
3.3 Содержание дисциплины 9
3.4 Тематика семинарских/практических и лабораторных занятий 11
3.5 Тематика курсовых проектов (курсовых работ) 17
4 Учебно-методическое и информационное обеспечение 17
4.1 Нормативные документы и ГОСТы 17
4.2 Основная литература 17
4.3 Дополнительная литература 17
4.4 Электронные образовательные ресурсы 17
4.5 Лицензионное и свободно распространяемое программное обеспечение 17
4.6 Современные профессиональные базы данных и информационные справочные системы 18
5 Материально-техническое обеспечение 18
6 Методические рекомендации 18
6.1 Методические рекомендации для преподавателя по организации обучения 18
6.2 Методические указания для обучающихся по освоению дисциплины 18
7 Фонд оценочных средств 19
7.1 Методы контроля и оценивания результатов обучения 19
7.2 Шкала и критерии оценивания результатов обучения 19
7.3 Оценочные средства 29

1 Цели, задачи и планируемые результаты обучения по дисциплине

К **основным целям** освоения дисциплины относится:

- формирование теоретических знаний о современных принципах, методах и средствах анализа данных, практических умений и навыков по применению современных методов анализа данных в различных сферах человеческой деятельности
- закрепление получаемых в семестре знаний и навыков на практике;
- формирование взаимосвязей, получаемых в семестре знаний и навыков с изученными ранее и изучаемых параллельно с данной дисциплиной;
- подготовка студентов к деятельности в соответствии с квалификационной характеристикой бакалавра.

К **основным задачам** дисциплины относятся:

- изучение существующих технологий подготовки данных к анализу
- изучение методов статистического анализа данных
- изучение и освоение теоретического материала, как в процессе контактной, так и в ходе самостоятельной работы;
- выполнение предоставленных практических заданий различных форм, как в процессе контактной, так и в ходе самостоятельной работы;

самостоятельная работа над тематикой дисциплины для формирования компетенций основной образовательной программы (далее, ООП)..

Обучение по дисциплине «Технологии анализа данных» направлено на формирование у обучающихся следующих компетенций:

Код и наименование компетенций	Индикаторы достижения компетенции
УК-6. Способен определять и реализовывать приоритеты собственной деятельности и способы ее совершенствования на основе самооценк	**знать:** методики самооценки, самоконтроля и саморазвития с использованием подходов здоровьесбережения; **уметь:** решать задачи личностного и профессионального развития, определять и реализовывать приоритеты совершенствования собственной деятельности; применять методики самооценки и самоконтроля; применять методики, позволяющие улучшить и сохранить здоровье в процессе жизнедеятельности; **владеть:** технологиями и навыками управления своей познавательной деятельностью и ее совершенствования на основе самооценки, самоконтроля и принципов самообразования в течение всей жизни, в том числе с использованием здоровьесберегающих подходов и методик.
ПК-4. Способен к разработке систем управления базами данных	***Знать:*** теория баз данных, основные структуры данных, основные модели данных и их организация, принципы построения языков запросов и манипулирования

	данными, методы обработки данных, основы современных систем управления базами данных, методы поддержки, контроля и оптимизации баз данных, системы хранения и анализа баз данных, методы повышения надежности работы системы управления базами данных, методы построения баз знаний и принципы построения экспертных систем, синтаксис, особенности программирования и стандартные библиотеки выбранного языка программирования, конструкции распределенного и параллельного программирования, способы и механизмы управления данными, принципы организации, состав и схемы работы операционных систем, принципы управления ресурсами, методы организации файловых систем, принципы построения сетевого взаимодействия, основы архитектуры, устройства и функционирования вычислительных систем, архитектура и принципы функционирования коммуникационного оборудования, устройство и принципы функционирования информационных систем, стандарты информационного взаимодействия систем, принципы организации инфокоммуникационных систем, основы информационной безопасности, подходы к автоматизации и стандарты автоматизации организации, локальные правовые акты, действующие в организации, английский язык на уровне чтения технической документации в области информационных и компьютерных технологий ***Уметь:*** Идентифицировать класс разрабатываемой системы управления базами данных в зависимости от выполняемых ею задач, определенных в техническом задании на разработку системы управления базами данных, идентифицировать класс разрабатываемой системы управления базами данных в зависимости от аппаратных средств, определенных в техническом задании на разработку системы управления базами данных, создавать блок-схемы алгоритмов функционирования разрабатываемых

	компонентов системы управления базами данных, оценивать вычислительную сложность алгоритмов функционирования разрабатываемых компонентов системы управления базами данных, применять языки программирования, определенные в техническом задании на разработку системы управления базами данных, для написания программного кода, осуществлять подготовку и сохранение резервных копий данных, применять нормативно-техническую документацию при использовании систем управления базами данных ***Владеть:*** навыками получения технической документации на разработку системы управления базами данных, изучения технической документации на разработку системы управления базами данных, разработки структуры системы управления базами данных в целом и ее отдельных компонентов, создания блок-схемы системы управления базами данных, разработки системы многозадачного и многопользовательского режимов, разработки системы администрирования данных, разработки системы поддержки транзакционных механизмов, разработки системы масштабируемости системы управления базами данных, разработки системы контроля целостности данных, разработки системы безопасности системы управления базами данных, разработки системы резервного копирования, написанияисходного кода системы управления базами данных на языке программирования системы управления базами данных, передачи исходного кода системы управления базами данных на тестирование
ПК-6. Способен к интеграции разработанного системного программного обеспечения	**Знать:** Основные серверы интеграции, их основные возможности и особенности, скриптовые языки, средства управления заданиями, механизмы мониторинга системы управления базами данных, основы систем управления базами данных, принципы организации, состав и схемы работы операционных систем, принципы построения сетевого взаимодействия, основные методы разработки

	программного обеспечения, основы архитектуры, устройства и функционирования вычислительных систем, архитектура и принципы функционирования коммуникационного оборудования, устройство и принципы функционирования информационных систем, методики тестирования разрабатываемых информационных систем, стандарты информационного взаимодействия систем, программные средства и платформы инфраструктуры информационных технологий, аппаратные средства и платформы инфраструктуры информационных технологий, основы информационной безопасности, теория системного анализ **Уметь:** устанавливать и настраивать серверы интеграции, налаживать автоматическую сборку разработанного системного программного обеспечения, определять порядок сборки разработанного системного программного обеспечения с учетом зависимостей в компонентахисать скрипты автоматизации сборки на скриптовых языках, писать задания для сред управления заданиями, работать в используемой системе управления требованиями, работать в используемой системе управления версиями **Владеть:** навыками планирования архитектуры инфокоммуникационной системы, планирования использования аппаратных и программных средств, выбора стратегии интеграции и практикуемых способов сборки разработанного системного программного обеспечения, определения порядка управления версиями сборок разработанного системного программного обеспечения, подготовки интеграционного сервера и настройка автоматической сборки разработанного системного программного обеспечения

2 Место дисциплины в структуре образовательной программы

Дисциплина относится к числу учебных дисциплин обязательной части.

Дисциплина взаимосвязана логически и содержательно-методически со следующими дисциплинами и практиками ООП:

. Большие данные и теория вычислительного эксперимента

3 Структура и содержание дисциплины

Общая трудоемкость дисциплины составляет 4 зачетных единицы, т.е. 144 академических часов, 6 часов лекций, 30 часов лабораторные занятия и 108 часов – самостоятельная работа студентов, экзамен - 4 семестр.

3.1 Виды учебной работы и трудоемкость (по формам обучения)

3.1.1 Очная форма обучения

№ п/п	Вид учебной работы	Количество часов	Семестры	
			4	
1	**Аудиторные занятия**	**36**		
	В том числе:			
1.1	Лекции	6	6	
1.2	Семинарские/практические занятия			
1.3	Лабораторные занятия	30	30	
2	**Самостоятельная работа**	**108**		
	В том числе:			
2.1	Подготовка к практическим занятиям	108	108	
3	**Промежуточная аттестация**			
	Зачет/диф.зачет/экзамен		экзамен	
	Итого:	**144**		

3.2 Тематический план изучения дисциплины (по формам обучения)

3.2.1 Очная форма обучения

№ п/п	Разделы/темы дисциплины	Трудоемкость, час					
		Всего	Аудиторная работа				Самостоятельная работа
			Лекции	Семинарские/ практические занятия	Лабораторные занятия	Практическая подготовка	
1.1	Тема 1. Предмет, задачи и содержание дисциплины	11	1		2		8
1.2	Тема 2. Постановка задачи в анализе данных	10			2		8
1.3	Тема 3. Основы проверки статистических гипотез	11	1		2		8
1.4	Тема 4. Анализ одной и двух нормальных выборок	9			2		7
1.5	Тема 5. Однофакторный анализ	10	1		2		7
1.6	Тема 6. Дисперсионный анализ	9			2		7
1.7	Тема 7. Двухфакторный анализ	10	1		2		7
1.8	Тема 8. Линейный и нелинейный регрессионный анализ	9			2		7
1.9	Тема 9. О проверке предпосылок в задаче регрессионного анализа	10	1		2		7
1.10	Тема 10. Независимость признаков	9			2		7
1.11	Тема 11. Критерий согласия	10	1		2		7
1.12	Тема 12. Временные ряды	9			2		7
1.13	Тема 13. Анализ временных рядов на компьютере	10	1		2		7
1.14	Тема 14. Многомерный анализ и другие статистические методы	8			1		7
1.15	Тема 15. Кластерный анализ	9	1		1		7
Итого		**144**	**6**		**30**		**108**

3.3 Содержание дисциплины

Предмет, задачи и содержание дисциплины.

Содержание: Роль анализа данных в теории познания. Многообразие анализа данных. Основные этапы развития методов и средств анализа данных. Структура курса, его место и роль в подготовке бакалавра, связь с другими дисциплинами и ознакомление с компетенциями, которыми должны обладать студенты после изучения данной дисциплины.

Постановка задачи в анализе данных. Содержание:Формула для вероятностей сумм независимых, одинаково распределенных, решетчатых случайных величин выраженная через сумму вероятностей полиномиального распределения.

Основы проверки статистических гипотез.

Содержание: Статистические модели. Проверка статистических гипотез. Схема испытаний Бернулли. Критерий знаков для одной выборки. Критерий Манна-Уитни. Критерий Уилкоксона. Анализ повторных парных наблюдений с помощью знаковых рангов.

Анализ одной и двух нормальных выборок.

Содержание: Об исследовании нормальных выборок. Одна выборка. Две выборки. Парные данные.

Однофакторный анализ.

Содержание: Постановка задачи. Непараметрические критерии проверки однородности. Критерий Краскела-Уоллиса. Критерий Джонкхиера. Практический пример.

Дисперсионный анализ.

Содержание: Оценивание эффектов обработки. Оценивание эффектов обработки в нормальной модели. Доверительные интервалы. Метод Шеффе множественных сравнений.

Двухфакторный анализ.

Содержание: Связь задач двухфакторного и однофакторного анализа.Таблица двухфакторного анализа. Непараметрические критерии проверки гипотезы об отсутствии эффектов обработки. Аддитивная модель данных двухфакторного эксперимента при независимом действии факторов.

Линейный и нелинейный регрессионный анализ.

Содержание: Модель линейного регрессионного анализа. О стратегии, методах и проблемах регрессионного анализа. Простая линейная регрессия.

О проверке предпосылок в задаче регрессионного анализа.

Содержание: Непараметрическая линейная регрессия. Практический пример. Регрессионный анализ в пакетах Excel, STADIA и SPSS.

Независимость признаков.

Содержание: Инструменты и стратегия исследования связи признаков. Связь признаков в количественных шкалах. Нормальная корреляция. Коэффициент корреляции.

Критерий согласия.

Содержание: Критерии согласия Колмогорова и омега-квадратв случае простой гипотезы. Критерий согласия хи-квадрат К.Пирсона для простой гипотезы. Критерии согласия для сложной гипотезы. Критерий согласия хи-квадрат Фишера для сложной гипотезы.

Временные ряды.

Содержание: Анализ временных рядов и его разделы. Цели, этапы и методы анализа временных рядов. Порядок анализа временных рядов.

Графические методы анализа временных рядов. . Методы исследования структуры стационарного временного ряда. Цели и методы анализа.

Анализ временных рядов на компьютере.

Содержание: Анализ временных рядов в Excel и SPSS. Обзор возможностей. Подбор тренда и прогнозирование. Устранение сезонной компоненты.

Многомерный анализ и другие статистические методы.

Содержание: Многомерный статистический анализ. Факторный анализ. Дискриминантный анализ.

Кластерный анализ.

Содержание: Многомерное шкалирование. Методы контроля качества. Использование статистических пакетов.

3.4 Тематика семинарских/практических и лабораторных занятий

ЛР-1	**Парная регрессия и корреляция**	6 ак. часов
Цель выполнения лабораторной работы: Ознакомление с парной регрессией и корреляцией, основными критериями качества регрессионных моделей		
Результат: Построение модели парной линейной регрессии реальной экономической задачи и оценка адекватности и качества этой модели		
Порядок выполнения лабораторной работы: 1.Для каждого двумерного набора данных, приведенных в Индивидуальном задании: 1.1. Рассчитайте коэффициент корреляции. Сделайте вывод о силе и направлении связи между переменными. 1.2. Постройте поле корреляции. Правильно ли сделан вывод о взаимосвязи между переменными в п. 1.1? 2. Определите, какая из переменных является фактором (причиной), а какая - результирующим показателем (следствием)? 3.Постройте уравнение линейной регрессии методом наименьших квадратов, воспользовавшись формулами определения коэффициентов регрессии, приведенными в лекции. Дать экономическую интерпретацию коэффициента регрессии. 4.Рассчитайте прогнозируемые значения результирующего показателя $\hat{Y}$ для каждого значения фактора Xj на основании построенного уравнения регрессии. Постройте новое поле корреляции и нанесите на него линию регрессии, используя рассчитанные значения $\hat{Y}$. 5.Определите стандартную ошибку регрессии и коэффициент детерминации. Сделайте вывод, какая доля вариации результирующего показателя У объясняется с помощью уравнения регрессии. 6. Используйте функцию ЛИНЕЙН для определения коэффициентов регрессии и других показателей регрессии. Проверьте ранее выполненные расчеты. 7.Осуществить проверку значимости параметров уравнения регрессии с помощью t-критерия Стьюдента *(a* = 0,05). 8. Вычислить коэффициент детерминации, проверить значимость уравнения регрессии с помощью *F* - критерия Фишера *(a* = 0,05), найти среднюю ошибку аппроксимации. Сделать вывод о качестве модели. 9.Защита лабораторной работы.		
Контрольные вопросы: 1. Какими способами можно обнаружить наличие взаимосвязи между двумя переменными? Кратко охарактеризуйте каждый из них. 2. Как рассчитывается коэффициент корреляции? Как можно интерпретировать его значения? 3. Какие проблемы, содержащиеся в данных, могут привести в неверной интерпретации значения коэффициента корреляции? 4. Что представляет собой математическая модель парной линейной регрессии? Как называются переменные, входящие в ее состав?		

5. Дайте определение функции регрессии. Чем отличаются функция регрессии и регрессионная модель? Покажите на графике индивидуальное значение y(x).
6. Назовите причины наличия в регрессионной модели случайного отклонения. Покажите случайное отклонение на графике.
7. Какова основная задача эконометрики и регрессионного анализа? Назовите его основные этапы и кратко охарактеризуйте каждый.
8. В чем суть метода наименьших квадратов? Запишите уравнение парной линейной регрессии. Как рассчитывается прогнозируемое значение результирующего показателя? Покажите его на графике.
9. Как рассчитываются коэффициенты регрессии с помощью метода наименьших квадратов? Что показывает каждый из этих коэффициентов?
10. Какие характеристики 'используются для оценки точности предсказаний по уравнению регрессии? Опишите каждую из них.
11. Для чего используется функция ЛИНЕЙН? ЧТО определяют ее аргументы «константа» и «статистика»?

Кроме того, на защите лабораторной работы студент должен по указанию преподавателя выполнить следующие действия:

1.Указать в массиве результатов функции ЛИНЕЙН значения известных ему показателей.

2.Дать интерпретацию заданного значения коэффициента корреляции.

3.Рассчитать на основании заданной функции регрессии прогнозируемое значение результирующего показателя, соответствующее заданному значению фактора.

4.Дать интерпретацию заданного значения коэффициента детерминации.

ЛР-2	**Регрессионный и дисперсионный анализ парной линейной регрессии**	6 ак. часов

Цель выполнения лабораторной работы: Изучение предпосылок МНК , критериев оценок параметров парной линейной регрессии и прогнозирование с помощью данной модели.

Результат: Модель парной линейной регрессии, которую можно использовать для прогнозирования.

Порядок выполнения лабораторной работы:

1. Выполнить лабораторную работу, используя опцию «анализ данных»
2. Используя схему дисперсионного анализа, вычислить остатки; найти остаточную сумму квадратов; факторную сумму квадратов и общую сумму квадратов, оценить дисперсию остатков*;* построить графики остатков.
3. Проверить выполнение предпосылок МНК.
4. Осуществить прогнозирование среднего значения показателя *У* при уровне значимости $a = 0{,}05$ и $a = 0{,}1$, если прогнозное значения фактора X составит 80% от его максимального значения.
5. Осуществить проверку значимости параметров уравнения регрессии с помощью t-критерия Стьюдента *(*$a = 0{,}05$*) и (*$a = 0{,}1$*).*

6. Вычислить коэффициент детерминации и дать его экономическую интерпретацию, проверить значимость уравнения регрессии с помощью *F* - критерия Фишера *(a* = 0,05) и *(a* = 0,1), найти среднюю ошибку аппроксимации. Сделать вывод о качестве модели.

7. Защита лабораторной работы.

Контрольные вопросы:

1.Сформулировать основную идею дисперсионного анализа.

2.Изобразить схему дисперсионного анализа.

3. Сформулируйте пять предпосылок МНК.

4. Что такое дисперсия?

5. Что такое математическое ожидание?

6. Что такое ковариация?

7.Сформулировать основные свойства оценок.

8.Что такое линейный коэффициент корреляции?

9.Что такое линейный коэффициент детерминации?

10. Для чего нужны расчеты доверительных интервалов?

11. В анализе данных в опции «регрессия» знать все названия показателей и их значение.

ЛР-3	**Построение модели множественной регрессии и корреляции.**	6 ак. часа

Цель выполнения лабораторной работы: Изучение принципов построения уравнения множественной регрессии на основе метода наименьших квадратов с целью дальнейшего анализ и прогноза.

Результат: Модель множественной регрессии с этапом идентификации и верификации, оценка качества и точности модели для дальнейшего долгосрочного прогнозирования.

Порядок выполнения лабораторной работы:

1. Постройте уравнение множественной регрессии методом МНК.
2. Постройте уравнение множественной линейной регрессии для данных, приведенных в Индивидуальном задании, используя функцию ЛИНЕЙН. Запишите построенное уравнение.
3. Дайте словесную интерпретацию каждому из коэффициентов регрессии.
4. Рассчитайте прогнозируемое значение результирующего показателя при значениях факторов, указанных в Индивидуальном задании.
5. Определите стандартную ошибку регрессии и коэффициент детерминации. Сделайте выводы об адекватности уравнения регрессии эмпирическим данным.
6. Проверьте значимость коэффициента детерминации. Сделайте вывод об адекватности уравнения регрессии данным генеральной совокупности.
7. Определите стандартные ошибки каждого из коэффициентов регрессии.
8. Проверьте значимость каждого из коэффициентов регрессии любым способом (не менее 2-х способов).
9. Определите, какой из факторов сильнее влияет на результирующий показатель (тремя способами).
10. Совпадает ли ваш вывод с результатами, полученными в п.7 и 8?

11. Выполните проверку рассчитанных вами показателей с помощью режима «Регрессия».
12. Постройте диагностическую диаграмму. Сделайте вывод.
13. Обобщите выводы о качестве построенного уравнения регрессии: а) о точности прогнозов, сделанных на основе уравнения регрессии; б) об адекватности уравнения данным генеральной совокупности; в) о необходимости исключения факторов из уравнения регрессии; г) о других проблемах и способах их решения.
14. Если необходимо исключить факторы из уравнения регрессии, запишите вид нового уравнения, которое вы предлагаете построить.
15. Выполните все задания из индивидуального варианта.
16. Защита лабораторной работы.

Контрольные вопросы:

1.Запишите теоретическое и эмпирическое уравнение множественной линейной регрессии.

В чем отличие между ними?

2.Какое соотношение должно выполняться между количеством переменных в линейной регресс ионной модели и числом наблюдений? Поясните свой ответ на примере парной линейной регрессии.

3.В чем суть МНК (с выводом). Запишите систему линейных нормальных уравнений для определения коэффициентов регрессии. Что показывают коэффициенты множественной регрессии?

4.Какие характеристики используются для оценки точности предсказаний по уравнению множественной регрессии? Опишите каждую из них.

5.Какие условия относительно генеральной совокупности должны выполняться, чтобы

статистические выводы о качестве уравнения регрессии были обоснованными?

6.Поясните понятия гомоскедастичности и гетероскедастичности.

7.Поясните понятие автокорреляции остатков.

8.Опишите механизм проверки гипотезы о значимости коэффициента детерминации.

Какова цель проверки этой гипотезы? Как она взаимосвязана с проверкой значимости коэффициентов регрессии?

9.Что показывают стандартные ошибки коэффициентов регрессии? Для чего они используются? Как можно проверить значимость коэффициентов регрессии?

1 О. Как можно сравнить влияние различных факторов на результирующий показатель?

Приведите формулы расчета соответствующих показателей.

11.Для чего строится диагностическая диаграмма? Как может интерпретироваться ее структура?

Кроме того, студент должен

1.Объяснить значение любого показателя в массиве результатов функции ЛИНЕЙН.

2.Рассчитать на основании заданной функции множественной регрессии прогнозируемое значение результирующего показателя, соответствующее заданным значениям факторов.

3.Указать в массиве результатов режима «Регрессия» значения известных ему показателей.

ПРИМЕЧАНИЕ: студент выполняет все пункты общего и индивидуального задания.

ЛР-4	**Построение и сравнительный анализ различных видов моделей нелинейной парной регрессии.**	5 ак. часов

Цель выполнения лабораторной работы: Изучение преимуществ и особенностей использования нелинейных моделей парной регрессии и корреляции.

Результат: Выбор лучшей по качеству и адекватности модели нелинейной регрессии по основным критериям.

Порядок выполнения лабораторной работы:

1.На основе заданных значений переменных (см. Индивидуальное задание) постройте поле корреляции и выполните спецификацию регрессионной модели.

2.Постройте линейное уравнение парной регрессии методом МНК, функцией «ЛИНЕЙН» и, используя режим «Регрессия», выполните проверку его качества:

2.1. Сделайте вывод об адекватности уравнения эмпирическим данным.

2.2. Проверьте адекватность уравнения данным генеральной совокупности любым из способов.

2.3. Проверьте с помощью диагностической диаграммы, соблюдаются ли предпосылки МНК

3.Постройте *первое нелинейное уравнение – «СТЕПЕННАЯ ФУНКЦИЯ»*

3.1. Подготовьте исходные данные для определения коэффициентов уравнения.

3.2. Постройте нелинейное уравнение регрессии МНК и, используя режим «Регрессия». Запишите уравнение.

3.3. Сделайте выводы о качестве уравнения регрессии согласно п.2.

4.Постройте *второе нелинейное уравнение – «ПОКАЗАТЕЛЬНАЯ ФУНКЦИЯ»*

4.1. Подготовьте исходные данные для определения коэффициентов уравнения.

4.2. Постройте нелинейное уравнение регрессии МНК и, используя режим «Регрессия». Запишите уравнение в преобразованном и исходном виде.

4.3. Сделайте выводы о качестве уравнения регрессии согласно п.2.

*5.*Постройте *полиномиальное* уравнение регрессии.

5.1. Подготовьте исходные данные для определения коэффициентов уравнения.

5.2. Постройте полиномиальное уравнение регрессии, используя режим «Регрессия». Запишите уравнение.

5.3. Сделайте выводы о качестве уравнения регрессии. Кроме проверок, перечисленных в п.2, выполните также проверку значимости угловых коэффициентов полиномиального уравнения.

6.Постройте поле корреляции, нанесите на него линии тренда, соответствующие построенным функциям регрессии, с указанием уравнения каждой функции и коэффициента детерминации. Сравните результаты с уравнениями, построенными в предыдущих пунктах.

7.Сравните качество построенных регрессионных уравнений. Выберите наилучшее уравнение и рассчитайте на его основе прогнозируемое значение результирующего показателя при значении фактора, указанном в Индивидуальном задании.

8.Для указанных моделей найти коэффициенты детерминации, коэффициенты эластичности и средние относительные ошибки аппроксимации. Сравнить модели по этим характеристикам и сделать вывод (в виде сводной таблицы данных).

9.Защита лабораторной работы.

Контрольные вопросы:

1.Какими способами осуществляется спецификация регрессионной модели? Назовите основные критерии качества регрессионных моделей.

2. Каким образом можно найти коэффициенты эмпирического уравнения степенной функции? Приведите примеры использования этой модели.

3.Каким образом можно найти коэффициенты эмпирического уравнения логарифмической функции? Приведите примеры использования этой модели.

4.Каким образом можно найти коэффициенты эмпирического уравнения гиперболической функции? Приведите примеры использования этой модели.

5.Каким образом можно найти коэффициенты эмпирического уравнения показательной функции? Приведите примеры использования этой модели.

6.Каким образом можно найти коэффициенты эмпирического уравнения экспоненциальной функции? Приведите примеры использования этой модели.

7.Каким образом можно найти коэффициенты эмпирического уравнения полиномиальной функции второго порядка? Приведите примеры использования этой модели.

8.Каким образом можно найти коэффициенты эмпирического уравнения полиномиальной функции третьего порядка? Приведите примеры использования этой модели.

9.Назовите критерии, по которым сравнивается качество регрессионных моделей. Какие ошибки спецификации можно обнаружить при анализе уравнений регрессии?

ЛР-5	**Построение моделей временного ряда**	5 ак. Часов

Цель выполнения лабораторной работы: Освоение навыков построения динамического временного ряда методом скользящего среднего и построение аддитивной и мультипликативной модели временного ряда.

Результат: Построенная модель временного ряда для долгосрочного прогнозирования

Порядок выполнения лабораторной работы:

1. Построить автокорреляционную функцию и сделать вывод о наличии сезонных колебаний.
2. Построить аддитивную модель временного ряда (для нечетных вариантов) или мультипликативную модель временного ряда (для четных вариантов).
3. Сделать прогноз на 2 квартала вперед.
4. Защита лабораторной работы.

Контрольные вопросы:

1. Общие понятия о системах эконометрических уравнений.
2. Структурная и приведенная формы модели.
3. Проблема идентификации. Необходимое условие идентифицируемости.
4. Проблема идентификации. Достаточное условие идентифицируемости.
5. Методы оценки параметров структурной формы модели.
6. Основные элементы временного ряда.
7. Автокорреляция уровней временного ряда и выявление его структуры.
8. Моделирование сезонных колебаний: аддитивная модель временного ряда.

9. Моделирование сезонных колебаний: мультипликативная модель временного ряда.
10. Критерий Дарбина-Уотсона.

.

3.5 Тематика курсовых проектов (курсовых работ)

Не предусмотрено

4 Учебно-методическое и информационное обеспечение

4.1 Нормативные документы и ГОСТы

1. Федеральный государственный образовательный стандарт высшего образования (уровень магистратуры) по направлению подготовки 01.04.02 Прикладная математика и информатика, утвержденный приказом Министерства науки и высшего образования Российской Федерации от 10.01.2018 №13.

4.2 Основная литература

1. Котиков, П. Е. Анализ данных : учебно-методическое пособие / П. Е. Котиков. — Санкт-Петербург : СПбГПМУ, 2019. — 48 с. — ISBN 978-5-907184-46-6. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/174498 (дата обращения: 27.09.2023). — Режим доступа: для авториз. пользователей.

2. Кадочникова, Е. И. Статистический анализ пространственных данных : учебное пособие / Е. И. Кадочникова, Ю. А. Варламова. — Казань : КФУ, 2023. — 140 с. — ISBN 978-5-00130-700-6. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/332354 (дата обращения: 27.09.2023). — Режим доступа: для авториз. пользователей.

4.3 Дополнительная литература

1. Дюк, В. А. Логический анализ данных : учебное пособие / В. А. Дюк. — Санкт-Петербург : Лань, 2020. — 80 с. — ISBN 978-5-8114-4180-8. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/126935 (дата обращения: 27.09.2023). — Режим доступа: для авториз. пользователей.

2. Кузьмин, В. И. Методы анализа данных : учебное пособие / В. И. Кузьмин, А. Ф. Гадзаов. — 2-е изд., перераб. и доп. — Москва : РТУ МИРЭА, 2020. — 155 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/171433 (дата обращения: 27.09.2023). — Режим доступа: для авториз. пользователей.

4.4 Электронные образовательные ресурсы

1. https://online.mospolytech.ru/course/view.php?id=6525 - МЕТОДЫ И ТЕХНОЛОГИИ АНАЛИЗА ДАННЫХ

4.5 Лицензионное и свободно распространяемое программное обеспечение

1. Microsoft Windows.
2. Microsoft Visio.
3. Microsoft Office.

4.6 Современные профессиональные базы данных и информационные справочные системы

Не предусмотрено

5 Материально-техническое обеспечение

Четыре компьютерных класса Ауд. АВ4805, АВ4809, АВ4810, АВ4811, оснащенные методическими материалами по дисциплине (лекции, практические задания).

.

6 Методические рекомендации

6.1 Методические рекомендации для преподавателя по организации обучения

1. При подготовке к занятиям следует предварительно проработать материал занятия, предусмотрев его подачу точно в отведенное для этого время занятия. Следует подготовить необходимые материалы – теоретические сведения, задачи и др. При проведении занятия следует контролировать подачу материала и решение заданий с учетом учебного времени, отведенного для занятия.

2. При проверке работ и отчетов следует учитывать не только правильность выполнения заданий, но и оптимальность выбранных методов решения, правильность выполнения всех его шагов.

3. При организации и проведения экзаменов в практико-ориентированной форме следует использовать утвержденные кафедрой Методические рекомендации.

6.2 Методические указания для обучающихся по освоению дисциплины

Изучение дисциплины осуществляется в строгом соответствии с целевой установкой в тесной взаимосвязи учебным планом. Основой теоретической подготовки студентов являются *аудиторные занятия, лабораторные работы.*

В процессе самостоятельной работы студенты закрепляют и углубляют знания, полученные во время аудиторных занятий, дорабатывают конспекты и записи, готовятся к проведению и обрабатывают результаты лабораторных работ, готовятся к промежуточной аттестации, а также самостоятельно изучают отдельные темы учебной программы.

На занятиях студентов, в том числе предполагающих практическую деятельность, осуществляется закрепление полученных, в том числе и в процессе самостоятельной работы, знаний. Особое внимание обращается на развитие умений и навыков установления связи положений теории с профессиональной деятельностью будущего специалиста.

Самостоятельная работа осуществляется индивидуально. Контроль самостоятельной работы организуется в двух формах:

- самоконтроль и самооценка студента;
- контроль со стороны преподавателей (текущий и промежуточный).

Критериями оценки результатов самостоятельной работы студента являются:

- уровень освоения студентом учебного материала;
- умения студента использовать теоретические знания при выполнении практических задач;
- сформированность компетенций;
- оформление материала в соответствии с требованиями.

.

7 Фонд оценочных средств

7.1 Методы контроля и оценивания результатов обучения

В процессе обучения используются следующие оценочные формы самостоятельной работы студентов, оценочные средства текущего контроля успеваемости и промежуточных аттестаций:

- выполнение лабораторных работ;
- экзамен.

7.2 Шкала и критерии оценивания результатов обучения

Показателем оценивания компетенций на различных этапах их формирования является достижение обучающимися планируемых результатов обучения по дисциплине.

Показатель	**Критерии оценивания**			
	2	**3**	**4**	**5**
УК-6. Способен определять и реализовывать приоритеты собственной деятельности и способы ее совершенствования на основе самооценки				
УК-6.1. Знать: методики самооценки, самоконтроля и саморазвития с использованием подходов здоровьесбережения УК-6.2. Уметь: решать задачи личностного и профессионального развития, определять и реализовывать приоритеты совершенствования собственной деятельности; применять методики самооценки и самоконтроля; применять методики, позволяющие улучшить и сохранить здоровье в процессе	Обучающийся демонстрирует полное отсутствие или недостаточное соответствие материалу дисциплины знаний, указанных в индикаторах компетенций дисциплины «Знать» (см. п. 3).	Обучающийся демонстрирует неполное соответствие следующих знаний, указанных в индикаторах компетенций дисциплины «Знать» (см. п. 3). Допускаются значительные ошибки, проявляется недостаточность знаний, по ряду показателей, обучающийся испытывает значительные затруднения при оперировании знаниями при их переносе на новые ситуации.	Обучающийся демонстрирует частичное соответствие следующих знаний, указанных в индикаторах компетенций дисциплины «Знать» (см. п. 3). Но допускаются незначительные ошибки, неточности, затруднения при аналитических операциях.	Обучающийся демонстрирует полное соответствие следующих знаний, указанных в индикаторах компетенций дисциплины «Знать» (см. п. 3). Свободно оперирует приобретенными знаниями.

жизнедеятельности УК-6.3. Владеть: технологиями и навыками управления своей познавательной деятельностью и ее совершенствования на основе самооценки, самоконтроля и принципов самообразования в течение всей жизни, в том числе с использованием здоровьесберегающих подходов и методик				
ПК-4. Способен к разработке систем управления базами данных				
ПК-4.1. Знать: теория баз данных, основные структуры данных, основные модели данных и их организация, принципы построения языков запросов и манипулирования данными, методы обработки данных, основы современных систем управления базами данных, методы поддержки, контроля и оптимизации баз данных, системы хранения и анализа баз данных, методы повышения	Обучающийся демонстрирует полное отсутствие или недостаточное соответствие материалу дисциплины знаний, указанных в индикаторах компетенций дисциплины «Знать» (см. п. 3).	Обучающийся демонстрирует неполное соответствие следующих знаний, указанных в индикаторах компетенций дисциплины «Знать» (см. п. 3). Допускаются значительные ошибки, проявляется недостаточность знаний, по ряду показателей, обучающийся испытывает значительные затруднения при оперировании знаниями при	Обучающийся демонстрирует частичное соответствие следующих знаний, указанных в индикаторах компетенций дисциплины «Знать» (см. п. 3). Но допускаются незначительные ошибки, неточности, затруднения при аналитических операциях.	Обучающийся демонстрирует полное соответствие следующих знаний, указанных в индикаторах компетенций дисциплины «Знать» (см. п. 3). Свободно оперирует приобретенными знаниями.

надежности работы системы управления базами данных, методы построения баз знаний и принципы построения экспертных систем, синтаксис, особенности программирования и стандартные библиотеки выбранного языка программирования, конструкции распределенного и параллельного программирования, способы и механизмы управления данными, принципы организации, состав и схемы работы операционных систем, принципы управления ресурсами, методы организации файловых систем, принципы построения сетевого взаимодействия, основы архитектуры, устройства и функционирования вычислительных систем, архитектура и принципы функционирования коммуникационного оборудования,		их переносе на новые ситуации.		

устройство и принципы функционирования информационных систем, стандарты информационного взаимодействия систем, принципы организации инфокоммуникационных систем, основы информационной безопасности, подходы к автоматизации и стандарты автоматизации организации, локальные правовые акты, действующие в организации, английский язык на уровне чтения технической документации в области информационных и компьютерных технологий ПК-4.2. Уметь: Идентифицировать класс разрабатываемой системы управления базами данных в зависимости от выполняемых ею задач, определенных в техническом задании на разработку системы управления базами данных, идентифицировать класс				

разрабатываемой системы управления базами данных в зависимости от аппаратных средств, определенных в техническом задании на разработку системы управления базами данных, создавать блок-схемы алгоритмов функционирования разрабатываемых компонентов системы управления базами данных, оценивать вычислительную сложность алгоритмов функционирования разрабатываемых компонентов системы управления базами данных, применять языки программирования, определенные в техническом задании на разработку системы управления базами данных, для написания программного кода, осуществлять подготовку и сохранение резервных копий данных, применять				

нормативно-техническую документацию при использовании систем управления базами данных ПК-4.3. Владеть: навыками получения технической документации на разработку системы управления базами данных, изучения технической документации на разработку системы управления базами данных, разработки структуры системы управления базами данных в целом и ее отдельных компонентов, создания блок-схемы системы управления базами данных, разработки системы многозадачного и многопользовательского режимов, разработки системы администрирования данных, разработки системы поддержки транзакционных механизмов, разработки				

системы масштабируемости системы управления базами данных, разработки системы контроля целостности данных, разработки системы безопасности системы управления базами данных, разработки системы резервного копирования, написанияисходного кода системы управления базами данных на языке программирования системы управления базами данных, передачи исходного кода системы управления базами данных на тестирование				
ПК-6. Способен к интеграции разработанного системного программного обеспечения				
ПК-6.1. Знать: Основные серверы интеграции, их основные возможности и особенности, скриптовые языки, средства управления заданиями, механизмы мониторинга системы управления базами данных,	Обучающийся демонстрирует полное отсутствие или недостаточное соответствие материалу дисциплины знаний, указанных в индикаторах компетенций дисциплины	Обучающийся демонстрирует неполное соответствие следующих знаний, указанных в индикаторах компетенций дисциплины «Знать» (см. п. 3). Допускаются значительные ошибки,	Обучающийся демонстрирует частичное соответствие следующих знаний, указанных в индикаторах компетенций дисциплины «Знать» (см. п. 3). Но допускаются незначительные ошибки,	Обучающийся демонстрирует полное соответствие следующих знаний, указанных в индикаторах компетенций дисциплины «Знать» (см. п. 3). Свободно оперирует приобретенными знаниями.

основы систем управления базами данных, принципы организации, состав и схемы работы операционных систем, принципы построения сетевого взаимодействия, основные методы разработки программного обеспечения, основы архитектуры, устройства и функционирования вычислительных систем, архитектура и принципы функционирования коммуникационного оборудования, устройство и принципы функционирования информационных систем, методики тестирования разрабатываемых информационных систем, стандарты информационного взаимодействия систем, программные средства и платформы инфраструктуры информационных технологий, аппаратные средства и платформы инфраструктуры информационных	«Знать» (см. п. 3).	проявляется недостаточность знаний, по ряду показателей, обучающийся испытывает значительные затруднения при оперировании знаниями при их переносе на новые ситуации.	неточности, затруднения при аналитических операциях.	

технологий, основы информационной безопасности, теория системного анализ ПК-6.2. Уметь: устанавливать и настраивать серверы интеграции, налаживать автоматическую сборку разработанного системного программного обеспечения, определять порядок сборки разработанного системного программного обеспечения с учетом зависимостей в компонентахисать скрипты автоматизации сборки на скриптовых языках, писать задания для сред управления заданиями, работать в используемой системе управления требованиями, работать в используемой системе управления версиями ПК-6.3. Владеть: навыками планирования архитектуры инфокоммуникаци				

онной системы, планирования использования аппаратных и программных средств, выбора стратегии интеграции и практикуемых способов сборки разработанного системного программного обеспечения, определения порядка управления версиями сборок разработанного системного программного обеспечения, подготовки интеграционного сервера и настройка автоматической сборки разработанного системного программного обеспечения				

Шкала оценивания	Описание
Отлично	Задание выполнено полностью и в срок. Отсутствуют ошибки в полученном результате. При процедуре защиты студент уверенно отвечает на контрольные вопросы, оперирует приобретенными знаниями и умениями, объясняет все этапы получения результата, его характеристики и причины их значений. Способен при необходимости доработать полученные результаты в соответствии с любыми незначительными изменениями в задании.
Хорошо	Задание выполнено полностью и в срок. Присутствуют незначительные ошибки в полученном результате. При процедуре защиты студент правильно отвечает на вопросы о ходе работы, оперирует приобретенными знаниями и умениями, однако возможны незначительные ошибки на дополнительные вопросы, в том числе и на вопросы для

	самоконтроля. Студент объясняет все этапы получения результата, его характеристики и причины их значений. Способен при необходимости доработать полученные результаты в соответствии с большинством незначительных изменений в задании.
Удовлетворительно	Задание выполнено либо со значительными ошибками, либо с опозданием. При процедуре защиты студент некорректно отвечает на некоторые дополнительные вопросы, в том числе и на вопросы для самоконтроля. Студент объясняет все этапы получения результата, его характеристики и причины их значений. Способен при необходимости доработать полученные результаты в соответствии с лишь некоторыми незначительными изменениями в задании.
Неудовлетворительно	Задание полностью не выполнено, либо выполнено не в срок и с грубыми ошибками. При процедуре защиты студент некорректно отвечает на большинство дополнительных вопросов, в том числе и на вопросы для самоконтроля. Не может объяснить этапы выполнения задания, характеристики и свойства полученного результата, причины и взаимосвязи между ними, исходными данными и своими действиями. Неспособен доработать полученные результаты в соответствии с незначительными изменениями в задании.

7.3 Оценочные средства

Типовые практические задания

1. Рассчитайте коэффициент корреляции. О чем свидетельствует его значение? Соответствует ли он полю корреляции?

2. Постройте уравнение линейной регрессии методом наименьших квадратов

3.Рассчитайте прогнозируемые значения результирующего показателя $\hat{Y}$для каждого значения фактора Xj на основании построенного уравнения регрессии. Постройте линию регрессии, добавив на поле корреляции рассчитанные значения $\hat{Y}$.

4.Определите стандартную ошибку регрессии и коэффициент детерминации. Какая доля вариации результирующего показателя У объясняется с помощью уравнения регрессии?Разработать программу, строящую гистограмму количества слов в тексте без учета регистра символов.

4. Постройте модель множественной регрессии. Обобщите выводы о качестве построенного уравнения регрессии: а) о точности прогнозов, сделанных на основе уравнения регрессии; б) об адекватности уравнения данным генеральной совокупности; в) о необходимости исключения факторов из уравнения регрессии; г) о других проблемах и способах их решения.