

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

**Федеральное государственное автономное
образовательное учреждение высшего образования
«Национальный исследовательский Нижегородский государственный университет
им. Н.И. Лобачевского»**

Институт экономики и предпринимательства
(факультет / институт / филиал)

УТВЕРЖДЕНО
решением ученого совета ННГУ
протокол от
«___» _____ 2022 г. № ___

Рабочая программа дисциплины

АНАЛИЗ ДАННЫХ

(наименование дисциплины (модуля))

Уровень высшего образования
бакалавриат

(бакалавриат / магистратура / специалитет)

Направление подготовки / специальность
38.03.05 Бизнес-информатика

(указывается код и наименование направления подготовки / специальности)

Направленность образовательной программы
Аналитические методы и информационные технологии поддержки принятия
решений в экономике и бизнесе

(указывается профиль / магистерская программа / специализация)

Форма обучения

Очная

(очная / очно-заочная / заочная)

Нижегород

2022 год

1. Место дисциплины в структуре ООП

Дисциплина *Б1.В.ДВ.02.01 Анализ данных* относится к части ООП направления подготовки 38.03.05 Бизнес-информатика, формируемой участниками образовательных отношений.

№ варианта	Место дисциплины в учебном плане образовательной программы	Стандартный текст для автоматического заполнения в конструкторе РПД
1	Блок 1. Дисциплины (модули) Часть, формируемая участниками образовательных отношений	Дисциплина Б1.В.ДВ.02.01 «Анализ данных» относится к части, формируемой участниками образовательных отношений ООП, направления подготовки 38.03.05 «Бизнес-информатика».

2. Планируемые результаты обучения по дисциплине, соотнесенные с планируемыми результатами освоения образовательной программы (компетенциями и индикаторами достижения компетенций)

Формируемые компетенции (код, содержание компетенции)	Планируемые результаты обучения по дисциплине (модулю), в соответствии с индикатором достижения компетенции		Наименование оценочного средства
	Индикатор достижения компетенции (код, содержание индикатора)	Результаты обучения по дисциплине	
<i>ПК-3.</i> Способен осуществлять разработку и исследование математических и компьютерных моделей поддержки принятия решений в экономике и бизнесе	<i>ПК-3.1.</i> Осуществляет разработку и исследование экономико-математических моделей.	<i>Знать</i> методы разработки и этапы исследования экономико-математических моделей <i>Уметь</i> разрабатывать и исследовать экономико-математические модели <i>Владеть</i> базовыми навыками по построению и анализу экономико-математических моделей	<i>Задания, тесты, собеседования, проект</i>
	<i>ПК-3.2.</i> Разрабатывает и применяет компьютерные модели в экономических исследованиях	<i>Знать</i> методы построения компьютерных моделей для экономических исследований <i>Уметь</i> разрабатывать и применять компьютерные модели в экономических исследованиях <i>Владеть</i> навыками работы в системах компьютерной математики для построения экономических моделей	<i>Задания, тесты, собеседования, проект</i>

3. Структура и содержание дисциплины

3.1 Трудоемкость дисциплины

	очная форма обучения	очно-заочная форма обучения	заочная форма обучения

Общая трудоемкость	5 ЗЕТ		
Часов по учебному плану	180		
в том числе			
аудиторные занятия (контактная работа):	59		
- занятия лекционного типа	28		
- занятия семинарского типа	28		
(практические занятия / лабораторные работы)			
самостоятельная работа	85		
КСР	3		
Промежуточная аттестация – экзамен/зачет	экзамен (36)		

3.2. Содержание дисциплины

Наименование и краткое содержание разделов и тем дисциплины	Всего (часы)			в том числе:												Самостоятельная работа обучающегося, часы		
				Контактная работа (работа во взаимодействии с преподавателем), часы														
				из них														
	Занятия лекционного типа	Занятия семинарского типа	Занятия лабораторного типа	Всего	Очная	Очно-заочная	Заочная	Очная	Очно-заочная	Заочная	Очная	Очно-заочная	Заочная	Очная	Очно-заочная	Заочная		
Очная					Очно-заочная	Заочная	Очная	Очно-заочная	Заочная	Очная	Очно-заочная	Заочная	Очная	Очно-заочная	Заочная	Очная	Очно-заочная	Заочная
Тема 1. Введение Анализ данных и его цели. Основы теории вероятностей: вероятности, их свойства, случайные величины, их характеристики, основные виды распределений и их свойства.	12			4			2						6			6		
Тема 2. Выборки и подготовка данных Типы данных, выборки, их виды и объемы, отличие от генеральной совокупности.	16			4			4						8			8		

ошибки, пропущенные и неопределенные данные, описательные статистики.																	
Тема 3. Графический анализ данных Основные виды графиков и сферы их применения.	8			2			2					4			4		
Тема 4. Проверка статистических гипотез. Статистические гипотезы. Критерии (критерий согласия)	8			2			2					4			4		
Тема 5. Поиск взаимосвязей и оценка их значимости Ковариация, корреляция, причинно-следственные связи, статистическая значимость.	8			2			2					4			4		
Тема 6. Линейная регрессия Метод наименьших квадратов, предпосылки, ограничения метода, прогноз.	16			4			4					8			8		
Тема 7. Сравнение выборок Межгрупповые сравнения, связанные и несвязанные выборки, критерии для сравнения	16			4			4					8			8		
Тема 8. Прикладные методы анализа данных Классификация данных. Кластерный анализ. Факторный анализ. Дискриминантный анализ.	24			6			6					12			12		

Тема 9. Применение проектно- ориентированн х методов обучения в изучении курса. Работа творческих коллективов над проектными работами. Защита проектных работ творческими коллективами	33			-		2					2			31		
Итого	141			28		28					56			85		

Семинарские занятия организуются, в том числе в форме практической подготовки, которая предусматривает участие обучающихся в выполнении отдельных элементов работ, связанных с будущей профессиональной деятельностью. Практическая подготовка предусматривает: выполнение проекта (учебно-исследовательской работы). На проведение семинарских занятий в форме практической подготовки отводится 3 часа

Практическая подготовка направлена на формирование и развитие:

- **практических навыков** в соответствии с профилем ОП (области профессиональной деятельности - 01 Образование и наука, 06 Связь и информационно-коммуникационные технологии, 07 Административно-управленческая и офисная деятельность, 08 Финансы и экономика):

аналитический:

- анализ запросов на изменение архитектуры и ИТ-инфраструктуры предприятия;
- анализ сферы деятельности, элементов архитектуры и ИТ-инфраструктуры предприятия;
- поиск инноваций в сфере ИКТ для решения задач в области профессиональной деятельности;
- анализ инноваций в экономике, управлении и информационно-коммуникативных технологиях;

научно-исследовательский:

- разработка и исследование экономико-математических моделей;
- разработка и применение компьютерных моделей в экономических исследованиях;
- подготовка научно-технических отчетов, презентаций, научных публикаций по результатам выполненных исследований;
- представление научного исследования в форме доклада;

технологический:

- применение современных языков программирования для разработки ИТ-сервисов предприятия;
- применение современных подходов и стандартов по управлению ИТ-сервисами;
- применение современных методологий и средств проектирования и построения архитектур электронного предприятия;
- моделирование и описание бизнес-процессов электронного предприятия в контексте его архитектуры;

инновационно-предпринимательский:

- выявление потребности в инновациях ИТ и исследование новых рынков;
- применение компьютерных программ и технологий при разработке бизнес-планов;
- поиск и анализ современные практик продвижения товаров и услуг;
- применение современных методов продвижения инновационных программно-информационных продуктов и услуг.

- **компетенций** - ПК-3 (Способен осуществлять разработку и исследование математических и компьютерных моделей поддержки принятия решений в экономике и бизнесе)

Текущий контроль успеваемости реализуется в рамках занятий семинарского типа.

4. Учебно-методическое обеспечение самостоятельной работы обучающихся

4.1. Виды самостоятельной работы по дисциплине:

- Подготовка к практическим занятиям
- Самостоятельное изучение некоторых теоретических аспектов теории экономического роста на основе работы с литературой.
- Подготовка к собеседованиям.
- Работа с литературой (аннотирование научных журнальных статей, посвященных теории экономического роста).
- Выполнение индивидуальных и коллективных заданий в рамках подготовки проектной работы.

4.2. Контрольные вопросы и задания для проведения текущего контроля и промежуточной аттестации по итогам освоения дисциплины приведены в п. 5.2.

Изучаемый курс считается освоенным, если по каждой из частей обучающимся продемонстрировано наличие определенного круга знаний, навыков, умений, позволяющих положительно оценить его работу по каждой части и, следовательно, по курсу в целом.

4.3. Для обеспечения самостоятельной работы обучающихся используется электронный курс Анализ данных (Капитанова О.В. (АнализД-БИ, <https://e-learning.unn.ru/course/view.php?id=4294>), созданный в системе электронного обучения ННГУ - <https://e-learning.unn.ru/>.

4.4. Методические указания по выполнению заданий для самостоятельной работы

Самостоятельная работа обучающихся – это планируемая работа, выполняемая по заданию и при методическом руководстве преподавателя, но без его непосредственного участия. Она способствует углублению и расширению знаний, формированию интереса к познавательной деятельности, овладению приемами процесса познания, развитию познавательных способностей.

Обучающиеся выполняют домашние задания, дают письменные ответы на вопросы, выполняют индивидуально и в группах задания и проводят необходимые действия по подготовке проектной работы, конспектируют научную и учебную литературу по изучаемым темам, готовят обзор публикаций по актуальным проблемам исследования по тематике проекта.

Качество самостоятельной работы обучающегося проверяется преподавателем во время практических занятий, при выполнении расчетно-графических работ, по результатам выполнения заданий, опросов, по результатам написания проектной работы и ее презентации, а также по степени активности участия во время занятий. По мере изучения дисциплины по составленным программным вопросам самим обучающимся осуществляется самоконтроль. Итоговый контроль представляет собой аттестацию обучающихся по всем видам работы.

4.5. Методические рекомендации по выполнению учебно-исследовательских (проектных) работ

Содержание этапа	Формируемые компетенции
1. Обоснование актуальности темы, её практической значимости	ПК-3

2. Теоретическая часть, включающая общетеоретический анализ проблемы, а также возможные применения методов математического и компьютерного моделирования	ПК-3
3. Практическая часть, включающая обоснование выбора методов математического и компьютерного моделирования изучаемой проблемы	ПК-3
4. Анализ результатов, выводы	ПК-3

В ходе изучения курса «Анализ данных» *предусматривается широкое использование проектно-ориентированных методов обучения* (в сочетании с внеаудиторной (самостоятельной) работой обучающихся).

В рамках этого метода обучающиеся (самостоятельно или в составе творческого коллектива) выполняют учебно-исследовательские работы. Приемлемые учебно-исследовательские работы представляют собой, как правило, работы следующего характера:

- ✓ строгий разбор, изучение и анализ статьи из списка литературы для чтения, или иной статьи, опубликованной в современной международной научной печати, относительно которой творческий коллектив думает, что она должна быть в списке литературы для чтения;

- ✓ углубленный анализ, обобщения, модификации статьи или цикла статей из списка литературы для чтения, или иных статей, опубликованных в современной международной научной печати, относительно которой творческий коллектив думает, что они представляют интерес для избранной темы исследовательской работы;

- ✓ поиск или разработка эффективных методов исследования, позволяющих дополнить результаты уже существующих опубликованных исследований.

Приветствуются и полностью оригинальные исследования по анализу данных. Однако использование интересной уже существующей работы как исходного пункта для проведения учебно-исследовательской работы может быть хорошим способом начать самостоятельные исследования.

Выполнение учебно-исследовательской работы (которая может носить характер учебно-научной или учебно-методической работы) строится по следующей схеме. Академическая группа подразделяется для выполнения текущей учебно-исследовательской работы на подгруппы по ~5 человек (подразделение группы на подгруппы выполняется преподавателем или же на основе добровольного объединения обучающихся в группы; возможны и иные способы, в том числе на основе случайного формирования состава группы).

Тема работы может быть сформулирована преподавателем или предложена подгруппой (творческим коллективом). Любой обучающийся, намеревающийся самостоятельно (в индивидуальном порядке) выполнить учебно-исследовательскую работу, должен сначала получить на это согласие лектора, а затем представить и обсудить с ним свой план работы (в течение первого месяца семестра).

В течение срока, отведенного на освоение курса, подгруппа разрабатывает тему учебно-исследовательского характера, подготавливает реферат по теме (лектору и на кафедру предоставляется окончательный текст работы с автографами авторов и электронный файл), и делает его презентацию (один или серия докладов на практических занятиях (семинарах)), на основе которой преподаватель определяет персональный вклад в общую работу каждого из членов подгруппы.

Требования к оформлению письменной работы (проекта): оптимальный объем творческой работы составляет в среднем 20-30 страниц машинописного текста. А4, Times New Roman, 14 пт, полуторный интервал (1,5 пт), выравнивание по ширине, нумерация страниц внизу от центра, номер 1 на титульном листе не ставится, красная строка – 1,25. Поля: левое – 30 мм, правое – 10 мм, верхнее – 20 мм, нижнее – 25 мм. Заголовки глав и параграфов отличаются по размеру и выделяются пустыми строками. Каждая глава начинается с новой страницы, после параграфа следует оставлять две пустых строки.

Крупные таблицы, рисунки и схемы выносятся в приложение. Подписи к рисункам располагаются под рисунком по центру; подписи к таблицам располагаются над таблицей по правому краю. Библиографические ссылки оформляются в соответствии с действующим стандартом.

Требования к содержанию проекта:

1. четкость и доступность изложения материала;
2. соответствие темы работы ее содержанию;
3. актуальность и практическая значимость работы;
4. эрудиция автора, умелое использование различных точек зрения по теме работы;
5. наличие собственных взглядов и выводов по проблеме;
6. умение использовать специальную терминологию и литературу по теме.

5. Фонд оценочных средств для промежуточной аттестации по дисциплине (модулю),

включающий:

5.1. Описание шкал оценивания результатов обучения по дисциплине

Уровень сформированности компетенций (индикатора достижения компетенций)	Шкала оценивания сформированности компетенций						
	плохо	неудовлетворительно	удовлетворительно	хорошо	очень хорошо	отлично	превосходно
	не зачтено		зачтено				
<u>Знания</u>	Отсутствие знаний теоретического материала. Невозможность оценить полноту знаний вследствие отказа обучающегося от ответа	Уровень знаний ниже минимальных требований. Имели место грубые ошибки.	Минимально допустимый уровень знаний. Допущено много негрубых ошибок.	Уровень знаний в объеме, соответствующем программе подготовки. Допущено несколько негрубых ошибок	Уровень знаний в объеме, соответствующем программе подготовки. Допущено несколько несущественных ошибок	Уровень знаний в объеме, соответствующем программе подготовки, без ошибок.	Уровень знаний в объеме, превышающем программу подготовки.
<u>Умения</u>	Отсутствие минимальных умений. Невозможность оценить наличие умений вследствие отказа обучающегося от ответа	При решении стандартных задач не продемонстрированы основные умения. Имели место грубые ошибки.	Продemonstrированы основные умения. Решены типовые задачи с негрубыми ошибками. Выполнены все задания, но не в полном объеме.	Продemonstrированы все основные умения. Решены все основные задачи с негрубыми ошибками. Выполнены все задания, в полном объеме, но некоторые с недочетами.	Продemonstrированы все основные умения. Решены все основные задачи. Выполнены все задания, в полном объеме, но некоторые с недочетами.	Продemonstrированы все основные умения, решены все основные задачи с отдельными несущественными недочетами, выполнены все задания в полном объеме.	Продemonstrированы все основные умения. Решены все основные задачи. Выполнены все задания, в полном объеме без недочетов
<u>Навыки</u>	Отсутствие владения материалом. Невозможность оценить наличие	При решении стандартных задач не продемонстрированы базовые	Имеется минимальный набор навыков для	Продemonstrированы базовые навыки	Продemonstrированы базовые навыки	Продemonstrированы навыки при решении нестандартн	Продemonstrирован творческий подход к решению нестандартн

	навыков вследствие отказа обучающегося от ответа	навыки. Имели место грубые ошибки.	решения стандартных задач с некоторыми недочетами	при решении стандартных задач с некоторыми недочетами	при решении стандартных задач без ошибок и недочетов.	ых задач без ошибок и недочетов.	ых задач
--	--	------------------------------------	---	---	---	----------------------------------	----------

Шкала оценки при промежуточной аттестации

Оценка		Уровень подготовки
	превосходно	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «превосходно», продемонстрированы знания, умения, владения по соответствующим компетенциям на уровне, выше предусмотренного программой
зачтено	отлично	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «отлично», при этом хотя бы одна компетенция сформирована на уровне «отлично»
	очень хорошо	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «очень хорошо», при этом хотя бы одна компетенция сформирована на уровне «очень хорошо»
	хорошо	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «хорошо», при этом хотя бы одна компетенция сформирована на уровне «хорошо»
	удовлетворительно	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «удовлетворительно», при этом хотя бы одна компетенция сформирована на уровне «удовлетворительно»
не зачтено	неудовлетворительно	Хотя бы одна компетенция сформирована на уровне «неудовлетворительно», ни одна из компетенций не сформирована на уровне «плохо»
	плохо	Хотя бы одна компетенция сформирована на уровне «плохо»

5.2. Типовые контрольные задания или иные материалы, необходимые для оценки результатов обучения.

5.2.1 Контрольные вопросы

№	Вопрос	Код компетенции (согласно РПД)
1.	Анализ данных и его цели.	ПК-3
2.	Основы теории вероятностей: вероятности и их свойства.	ПК-3
3.	Основы теории вероятностей: случайные величины и их характеристики.	ПК-3
4.	Основы теории вероятностей: основные виды распределений и их свойства.	ПК-3
5.	Типы данных	ПК-3
6.	Выборки, их виды и объемы	ПК-3
7.	Выборка и генеральная совокупность	ПК-3
8.	Ошибки, пропущенные и неопределенные данные	ПК-3
9.	Описательные статистики	ПК-3
10.	Основные виды графиков и сферы их применения.	ПК-3
11.	Статистические гипотезы.	ПК-3

12.	Критерий согласия	ПК-3
13.	Ковариация и корреляция.	ПК-3
14.	Причинно-следственные связи	ПК-3
15.	Статистическая значимость.	ПК-3
16.	Предпосылки и ограничения метода наименьших квадратов	ПК-3
17.	Линейная регрессия. Прогнозы	ПК-3
18.	Межгрупповые сравнения	ПК-3
19.	Связанные и несвязанные выборки	ПК-3
20.	Критерии для сравнения выборок	ПК-3
21.	Классификация данных.	ПК-3
22.	Кластерный анализ.	ПК-3
23.	Факторный анализ.	ПК-3
24.	Дискриминантный анализ.	ПК-3

5.2.2. Вопросы для собеседования для оценки компетенции ПК-3

1. Анализ данных и его цели.
2. Основы теории вероятностей: вероятности и их свойства.
3. Основы теории вероятностей: случайные величины и их характеристики.
4. Основы теории вероятностей: основные виды распределений и их свойства.
5. Типы данных
6. Выборки, их виды и объемы
7. Выборка и генеральная совокупность
8. Ошибки, пропущенные и неопределенные данные
9. Описательные статистики
10. Основные виды графиков и сферы их применения.
11. Статистические гипотезы.
12. Критерий согласия
13. Ковариация и корреляция.
14. Причинно-следственные связи
15. Статистическая значимость.
16. Предпосылки и ограничения метода наименьших квадратов
17. Линейная регрессия. Прогнозы
18. Межгрупповые сравнения
19. Связанные и несвязанные выборки
20. Критерии для сравнения выборок
21. Классификация данных.
22. Кластерный анализ.
23. Факторный анализ.
24. Дискриминантный анализ.

5.2.3. Тестовые вопросы для оценки компетенции ПК-3

1. Все сведения о чем-либо - это
 - ~данные
 - ~знания
 - =информация
 - ~аналитика
2. Данные, собранные вами для целей вашего исследования, называются
 - ~вторичными
 - ~качественными
 - ~исходными

- =первичными
- 3. Гистограмма, изображающая частоты встречаемости признака в убывающем порядке, на которую накладывается точечно-линейный график накопленных частот, отражающий частоту значений признака
 - =диаграмма Парето
 - ~гистограмма
 - ~огива
 - ~кумулята
- 4. Если имеющиеся данные описывают только часть какого-то класса явлений, и мы пытаемся по этой выборке охарактеризовать все целое, то говорят об
 - =аналитической статистике
 - ~описательной статистике
 - ~дескриптивной статистике
 - ~одномерной статистике
- 5. Значение признака, приходящееся на середину ранжированного ряда наблюдений
 - =медиана
 - ~мода
 - ~среднее
 - ~размах
- 6. Распределение с тремя значениями моды называется
 - =полимодальным
 - ~бимодальным
 - ~унимодальным
 - ~такого не может быть
- 7. Совокупность численных значений одного экономического показателя в разные моменты времени
 - =временной ряд
 - ~пространственные данные
 - ~временной срез
 - ~панельные данные
- 8. Сумма всех имеющихся наблюдений, деленная на их количество, называется
 - =среднее арифметическое
 - ~среднее геометрическое
 - ~среднее гармоническое
 - ~медиана
 - ~мода
- 9. Ураган Катрина был для статистики Нового Орлеана:
 - ~ошибкой третьего рода
 - =ошибкой второго рода
 - ~ошибкой первого рода
 - ~катастрофой
- 10. Формула Стерджеса основана на:
 - ~экспоненциальном распределении
 - ~распределении Парето
 - ~нормальном распределении
 - =биномиальном распределении
- 11. В интервале от -1 до 1 лежит
 - =коэффициент корреляции
 - ~коэффициент детерминации
 - ~коэффициент ковариации
 - ~коэффициент регрессии

12. Все возможные наблюдения, которые могут быть сделаны при данном комплексе условий, называются
- =генеральная совокупность
 - ~выборка
 - ~набор данных
 - ~ранжированный ряд
13. Средняя арифметическая квадрата отклонения наблюдаемых значений от средней арифметической - это
- =дисперсия
 - ~среднее квадратическое отклонение
 - ~ковариация
 - ~интерквартильный размах
14. Для получения достоверных результатов в анализе данных, выборка должна быть
- =репрезентативной
 - ~адекватной
 - ~правильной
 - ~достаточной
15. Для установления соответствия между размерностью величин применяется
- =нормирование данных
 - ~унификация данных
 - ~очистка данных
 - ~выравнивание данных
16. Для характеристики степени скошенности распределения данных используется
- =коэффициент асимметрии
 - ~коэффициент эксцесса
 - ~математическое ожидание
 - ~коэффициент Пирсона
17. Интерквартильный размах - это разница между
- =третьим и первым квартилем
 - ~третьим и вторым квартилем
 - ~вторым и первым квартилем
 - ~четвертым и вторым квартилем
18. Сумма произведений возможных значений дискретной случайной величины на соответствующие вероятности, называется
- =математическое ожидание
 - ~дисперсия
 - ~среднее квадратическое отклонение
 - ~центральный момент второго порядка
19. Если при любом объеме выборки n ее математическое ожидание равно оцениваемому параметру, то статистическая оценка этого параметра называется
- =несмещенной
 - ~эффективной
 - ~состоятельной
 - ~адекватной
20. Один из самых простых способов построения интервальной оценки – это применение неравенства
- =Чебышева
 - ~Лапласа
 - ~Стьюдента
 - ~Пирсона
21. Модели, которые описывают группы, на которые можно разделить объекты, данные о которых подвергаются анализу, называются...

- =модели кластеризации
 - ~итоговые модели
 - ~регрессионные модели
 - ~модели классификации
22. Представление правил в иерархической последовательной структуре - это построение
- =дерева решений
 - ~классификации
 - ~регрессии
 - ~кластеров
23. Квадратный корень присутствует в формуле расстояния:
- =Евклида
 - ~по Хеммингу
 - ~Чебышева
 - ~Махаланобиса
24. Определение значения зависимой переменной объекта (выбор из конечного множества) на основании значений других переменных, характеризующих данный объект, это
- =классификация
 - ~кластеризация
 - ~регрессия
 - ~факторизация
25. Разделение исследуемого множества объектов на группы «похожих» объектов - это задача ...
- =кластеризации
 - ~классификации
 - ~регрессии
 - ~диверсификации
26. Выберите среди перечисленных неиерархический алгоритм:
- =метод k-средних
 - ~метод ближайшего соседа
 - ~метод медиан
 - ~метод Уорда
27. Выберите среди перечисленных агломеративный алгоритм:
- =метод медиан
 - ~метод k-средних
 - ~алгоритм Fuzzy C-Means
 - ~кластеризация по Гюстафсону-Кесселю
28. Ковариационная матрица применяется для расчета расстояния:
- =Махаланобиса
 - ~по Хеммингу
 - ~Чебышева
 - ~Евклида

5.2.4. Задания для оценки компетенции ПК-3

Задание 1.

Для данных из таблицы¹ определите частоты, относительные (частоты) и накопленные частоты. Визуализируйте полученные результаты.

<i>Регион</i>	<i>Население</i>
Азия	4 641 054 775
Африка	1 340 598 147
Европа	747 636 026

¹ <https://www.worldometers.info/world-population/#region>

Латинская Америка и Карибы	653 962 331
Северная Америка	368 869 647
Австралия и Океания	42 677 813

Задание 2.

В таблице приведены данные по заработной плате работников малого предприятия. Визуализируйте данные и вычислите средние значения.

№ работника	Зарплата в тыс. руб.	№ работника	Зарплата в тыс. руб.	№ работника	Зарплата в тыс. руб.	№ работника	Зарплата в тыс. руб.
1	61,380	13	20,168	25	32,394	37	11,228
2	45,165	14	30,811	26	27,958	38	9,465
3	59,728	15	82,726	27	3,777	39	32,744
4	31,830	16	20,064	28	101,939	40	161,987
5	15,864	17	11,481	29	146,908	41	36,508
6	34,601	18	29,971	30	49,820	42	30,895
7	28,100	19	33,868	31	17,077	43	27,424
8	36,842	20	43,107	32	19,061	44	33,836
9	13,572	21	143,462	33	10,634	45	38,428
10	9,847	22	25,997	34	21,963	46	12,718
11	24,746	23	23,893	35	36,647	47	20,399
12	17,382	24	10,384	36	45,608	48	35,586

Задание 3.

Определите, какие данные нужно собрать, чтобы определить зависит ли успеваемость в вашей группе от пола студента. Соберите необходимые данные. Выполните предварительный анализ данных. Постройте необходимые графики и рассчитайте средние.

Задание 4.

Придумайте и проведите небольшой опрос (около 20-30 человек). Проведите предварительный анализ результатов. Составьте опрос так, чтобы у вас были и качественные, и количественные данные.

Задание 5.

По данным из Задания 2 определите математическое ожидание, дисперсию, среднее квадратическое отклонение, границы квартилей, интерквартильный размах, коэффициент вариации, коэффициенты асимметрии и эксцесса. Сделайте выводы. Нормируйте показатель, повторите расчеты, сравните результаты. Для расчетов используйте формулы из лекций, сравните результаты со встроенными функциями Excel. Постройте гистограмму, используя инструмент анализа данных. Сделайте выводы.

Задание 6.

Вычислите основные статистические характеристики для успеваемости вашей группы (Задание 3). Сделайте выводы.

Задание 7.

В файле Практика на Листе 1 (структура приведена в таблице ниже) содержится репрезентативная выборка доходов населения США в приведенных ценах. Что можно сказать об изменении доходов?

Person	1975	1985	1995	2005
--------	------	------	------	------

1	\$37 280,30	\$15 684,79	\$19 324,41	\$21 971,49
2	\$22 057,86	\$38 443,49	\$5 722,73	\$18 077,28
3	\$13 355,86	\$11 610,79	\$13 510,42	\$19 486,17
4	\$21 318,64	\$27 134,09	\$20 619,69	\$35 598,38
5	\$15 110,66	\$59 795,83	\$8 431,22	\$34 546,83
6	\$22 068,35	\$28 097,32	\$12 216,38	\$14 872,53

Задание 8.

4. В таблице ниже для нескольких слаборазвитых стран содержатся данные о детской смертности, уровне грамотности взрослого население, проценте учеников, закончивших начальную школу и ВВП на душу населения.

	Infant Mortality(deaths per thousand births)	%age adult literacy	%age finishing primary school	GNP per capita
Cuba	18	98	98	2000
Sri Lanka	20	85	92	3300
Costa Rica	19	94	84	5800
Vietnam	44	85	58	600
China	54	80	86	2400
South Africa	56	76	68	4000
Saudi Arabia	38	59	68	11000
Brazil	60	78	56	5600
Zimbabwe	68	82	76	1800
Morocco	68	42	76	3400
Pakistan	98	36	38	2100
Nigeria	86	44	56	1600

Исходя из этих данных создайте уравнение для прогнозирования детской смертности. Имеются ли в этом наборе выбросы? Проинтерпретируйте коэффициенты полученного уравнения. В пределах каких значений 95% прогнозов по детской смертности должны быть точными?

Задание 9.

С помощью кластерного анализа постройте распределение стран мира на группы (кластеры) по показателям, характеризующим благосостояния населения. Определите оптимальное количество кластеров.

Задание 10.

Постройте множественную линейную регрессию для зависимости ВВП на душу населения в странах мира от различных факторов, описывающих уровень благосостояния населения. Проверьте качество модели.

Задание 11.

С помощью кластерного анализа постройте распределение регионов России на группы (кластеры) по показателям, характеризующим уровень жилищных условий населения: структура жилищного фонда по формам собственности – частный (%), государственный (%), муниципальный (%), другие (%). Определите оптимальное количество кластеров.

Задание 12.

Проведите регрессионный анализ зависимости ВВП на душу населения в регионах ПФО от показателей, характеризующих численность занятого в экономике населения, индекс потребительских цен и среднюю заработную плату.

Задание 13.

Проведите кластерный анализ развития стран мира на основе показателей, описывающих состояние образования, состояние здравоохранения, качество жизни.

Задание 14.

Проведите факторный анализ данных на основе показателей, характеризующих финансово-экономическую деятельность компаний по видам экономической деятельности.

Задание 15.

С помощью кластерного анализа постройте распределение стран мира на группы (кластеры) по показателям, характеризующим благосостояния населения. Определите оптимальное количество кластеров.

Задание 16.

Проведите факторный анализ данных на основе показателей, описывающих образование в регионах РФ.

Задание 17.

Проведите кластерный анализ данных на основе показателей, характеризующих занятость экономически активного населения в регионах России.

5.2.5. Тематика проектных работ для оценки компетенции ПК-3

В рамках данной методики предполагается, что тему проектной работы формулирует сама группа.

- Тенденции развития нефтедобывающей отрасли в России.
- Исследования и разработки как важнейший фактор экономического роста.
- Анализ показателей благосостояния населения России.
- Развитие информационно-коммуникационных технологий в России.
- Демографическая ситуация в регионах России.
- Классификация недвижимости на рынке вторичного жилья в г. Нижний Новгород.
- Обработка данных социологических опросов.

6. Учебно-методическое и информационное обеспечение дисциплины

а) основная литература:

1. Анализ данных : учебник для вузов / В. С. Мхитарян [и др.] ; под редакцией В. С. Мхитаряна. — Москва : Издательство Юрайт, 2020. — 490 с. — (Высшее образование). — ISBN 978-5-534-00616-2. — Текст : электронный // ЭБС Юрайт [сайт]. — URL: <https://urait.ru/bcode/450166>.
2. Миркин, Б. Г. Введение в анализ данных : учебник и практикум / Б. Г. Миркин. — Москва : Издательство Юрайт, 2020. — 174 с. — (Высшее образование). — ISBN 978-5-9916-5009-0. — Текст : электронный // ЭБС Юрайт [сайт]. — URL: <https://urait.ru/bcode/450262>.

б) дополнительная литература:

1. Козлов, А. Ю. Статистический анализ данных в MS Excel : учебное пособие / А. Ю. Козлов, В. С. Мхитарян, В. Ф. Шишов. — Москва : ИНФРА-М, 2019. — 320 с. — (Высшее образование: Бакалавриат). - ISBN 978-5-16-004579-5. - Текст : электронный. - URL: <https://znanium.com/catalog/product/987337>.
2. Лонг, Д. Р. Книга рецептов: проверенные рецепты для статистики, анализа и визуализации данных : практическое руководство / Д. Лонг, П. Титор ; пер. с англ. Д. А. Беликова. - Москва : ДМК Пресс, 2020. - 510 с. - ISBN 978-5-97060-835-7. - Текст : электронный. - URL: <https://znanium.com/catalog/product/1210661>.

3. Сурина, Е. Е. Методы анализа экономической информации и данных : учебно-методическое пособие / Е. Е. Сурина. - 3-е изд., стер. - Москва : ФЛИНТА, 2020. - 130 с. - ISBN 978-5-9765-2499-6. - Текст : электронный. - URL: <https://znanium.com/catalog/product/1150895>.
4. Феррари, А. Анализ данных при помощи Microsoft Power BI и Power Pivot для Excel : практическое руководство / А. Феррари, М. Руссо ; пер. с англ. А. Ю. Гинько. - Москва : ДМК Пресс, 2020. - 288 с. - ISBN 978-5-97060-858-6. - Текст : электронный. - URL: <https://znanium.com/catalog/product/1210705>.
- в) программное обеспечение и Интернет-ресурсы (в соответствии с содержанием дисциплины)
 1. Операционная система Microsoft Windows
 2. Прикладное программное обеспечение Microsoft Office (Word, Excel, Power Point) и язык R с рабочей средой (например, R Studio)
 3. www.gks.ru / Федеральная служба государственной статистики
 4. <https://www.rug.nl/ggdc/productivity/pwt/?lang=en> / Penn World Table
 5. www.hks.harvard.edu/
 6. altaplana.com/Gate.resource.html
 7. data.worldbank.org/
 8. www.fira.ru
 9. statistika.ru
 10. www.cia.gov/library/publications/the-world-factbook/index.html
 11. data.un.org
 12. stats.oecd.org/index.aspx

7. Материально-техническое обеспечение дисциплины

Помещения представляют собой учебные аудитории для проведения учебных занятий, предусмотренных программой, оснащенные оборудованием и техническими средствами обучения: компьютерами, проектором или ЖК-телевизором, акустической системой и микрофоном (при необходимости), а также доской.

Помещения для самостоятельной работы обучающихся оснащены компьютерной техникой с возможностью подключения к сети "Интернет" и обеспечены доступом в электронную информационно-образовательную среду.

Программа составлена в соответствии с требованиями ФГОС ВО /ОС ННГУ по направлению подготовки 38.03.05 Бизнес-информатика, направленность «Аналитические методы и информационные технологии поддержки принятия решений в экономике и бизнесе».

Автор

доцент кафедры ММЭП, к.ф.-м.н. _____ Капитанова О.В.

Рецензент

доцент кафедры ДУМиЧА ИИТММ, к.ф.-м.н. _____ Круглов Е.В.

Заведующий кафедрой ММЭП

д.ф.-м.н., профессор _____ Кузнецов Ю.А.

Программа одобрена на заседании методической комиссии Института экономики и предпринимательства от « ____ » _____ 2022 года, протокол № .