

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

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

**Институт информационных технологий, математики и механики**

(факультет / институт / филиал)

УТВЕРЖДЕНО  
решением ученого совета ННГУ  
протокол от  
«30» ноября 2022 г. № 13

## **Рабочая программа дисциплины**

**Методы моделирования и прикладной анализ данных**

(наименование дисциплины (модуля))

Уровень высшего образования

**бакалавриат**

(бакалавриат / магистратура / специалитет)

Направление подготовки / специальность

**01.03.02 Прикладная математика и информатика**

(указывается код и наименование направления подготовки / специальности)

Направленность образовательной программы

**Прикладная математика и информатика (общий профиль)**

(указывается профиль / магистерская программа / специализация)

Форма обучения

**очная**

(очная / очно-заочная / заочная)

**Нижний Новгород**

**2023 г.**

## 1. Место дисциплины в структуре ООП

Дисциплина относится к части, формируемой участниками образовательных отношений.

Код дисциплины: Б1.В.ДВ.08.01 (Методы моделирования и прикладной анализ данных)

Осваивается в 7 семестре четвертого года обучения.

№ варианта	Место дисциплины в учебном плане образовательной программы	Стандартный текст для автоматического заполнения в конструкторе РПД
2	Блок 1. Дисциплины (модули) Часть, формируемая участниками образовательных отношений	Дисциплина Б1.В.ДВ.08.01 (Методы моделирования и прикладной анализ данных) относится к части ООП направления подготовки 01.03.02 «Прикладная математика и информатика», формируемой участниками образовательных отношений.

## 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	
<b>ПК-6.</b> Способен изучать и применять программное обеспечение, проводить расчётные работы и выполнять обработку результатов исследований	<b>ПК-6.1.</b> Знает методы применения современных программных комплексов, пакетов прикладных программ и автоматизированных систем для решения прикладных задач при проведении исследований.	<i>Знает математические модели прикладного анализа данных, их понятийный аппарат, основные утверждения, их доказательство (обоснование модели) и способ применения.</i>  <i>Знает типы и источники данных, теоретические основы разведочного анализа, кластерного анализа; метода главных компонент, факторного и регрессионного анализа, основы статистического вывода.</i>  <i>Знает постановки модельных задач для применения методов, понятийный аппарат, критерии качества построенных решений, способы представления решений в пакетах прикладных статистических программ, компьютерную диагностику построения решений.</i>	Собеседование, практическое задание. Тест
	<b>ПК-6.2.</b> Умеет самостоятельно проводить расчётные работы, выбирать и применять современные	<i>На предложенных комплектах данных умеет формализовать задачу, сформулировать цели исследования.</i>  <i>С помощью пакетов прикладных программ умеет применять разведочный анализ, кластерный анализ; метод</i>	

	программные комплексы, пакеты прикладных программ и автоматизированные системы, обрабатывать и анализировать полученные результаты	главных компонент, факторный анализ, проверять корректность применения методов.  Умеет формулировать результаты, обосновать применение метода, предложить интерпретацию.	
	<b>ПК-6.3.</b> Имеет практический опыт применения современного программного обеспечения для решения прикладных задач	Владеет навыками самостоятельного выполнения проектов с использованием пакетов статистических программ, включая выбор цели исследования, выбор метода, его теоретическое обоснование, анализ результатов, их структуризацию, интерпретацию, проверку качества решения, сравнительный анализ.	Практическое задание, проект

### 3. Структура и содержание дисциплины

#### 3.1. Трудоемкость дисциплины

	Очная форма обучения
<b>Общая трудоемкость</b>	<b>2 ЗЕТ</b>
<b>Часов по учебному плану</b>	<b>72</b>
<b>в том числе</b>	
<b>аудиторные занятия (контактная работа):</b>	<b>33</b>
- занятия лекционного типа	16
- занятия семинарского типа	16
- занятия лабораторного типа	
- текущий контроль (КСР)	1
<b>самостоятельная работа</b>	<b>39</b>
<b>Промежуточная аттестация – зачет</b>	

#### 3.2. Содержание дисциплины

Наименование и краткое содержание разделов и тем дисциплины	Всего (часы)	В том числе				
		Контактная работа (работа во взаимодействии с преподавателем), часы. Из них				Самостоятельная работа обучающегося, часы
		Занятия лекционного типа	Занятия семинарского типа	Занятия лабораторного типа	Всего	
Тема 1. Постановки задач. Источники и типы данных. Нормативно-правовые аспекты сбора данных. Интерпретация и выводы для управленческих решений. Примеры.	9	2	2		4	5

Тема 2. Методы разведочного, корреляционного и дисперсионного анализа. Введение в модели регрессии	9	2	2		4	5
Тема 3. Методы кластерного анализа. Теория и приложения.	9	2	2		4	5
Тема 4. Метод главных компонент. Теория и приложения.	14	4	4		8	6
Тема 5. Факторный анализ. Теория и приложения.	10	2	2		4	6
Тема 6. Основы статистического вывода для методов п. 2 – п. 5.	10	2	2		4	6
Тема 7. Применение методов для анализа модельных задач	10	2	2		4	6
Текущий контроль (КСР)	1				1	
Итого	<b>72</b>	<b>16</b>	<b>16</b>		<b>33</b>	<b>39</b>

Текущий контроль успеваемости реализуется в рамках занятий семинарского типа.  
Промежуточная аттестация проходит в традиционной форме (зачет).

#### **4. Учебно-методическое обеспечение самостоятельной работы обучающихся**

Самостоятельная работа студента при изучении дисциплины «Методы моделирования и прикладной анализ данных – 1» включает:

- ❖ выполнение практических заданий (Практикум)
- ❖ выполнение проектов (тематика в разделах Практикума)
- ❖ теоретическую проработку материала (практические задания теоретического характера)
- ❖ подготовку к собеседованию (зачету)
- ❖ тестирование

Практикум включает задания:

Стартовый практикум. Первичный анализ данных, анализ структуры и динамики (в EXEL)

Практикум 1. Первичный анализ данных. Метод К-средних (в SPSS®)

Практикум 2. Метод главных компонент (в SPSS®)

Практикум 3. Факторный анализ (в SPSS®)

Дополнительный практикум. Применение метода главных компонент для анализа данных, упорядоченных по времени (в SPSS®)

Практикум 4. Самостоятельный анализ данных (проект)

а также подготовку отчетов и презентаций с целью обсуждения результатов.

Содержание и порядок выполнения заданий, а также файлы данных приведены в [1].

Практические задания теоретического характера приведены в п. 5.2.2.

Для самоконтроля у студента имеется возможность удаленного тестирования по дистанционному лекционному курсу и выполнения заданий из УМК «Методы моделирования: анализ данных» ННГУ (<https://e-learning.unn.ru/course/enrol/index.php?id=827>, требуется регистрация).

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

**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						
	Не зачтено		Зачтено				
<u>Знания</u>	Отсутствие знаний теоретического материала.  Невозможно 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.	Продemonстрированы основные умения. Решены типовые задачи с негрубыми ошибками. Выполнены все задания, но не в полном объеме.	Продemonстрированы все основные умения. Решены все основные задачи с негрубыми ошибками. Выполнены все задания, в полном объеме, но некоторые с недочетами.	Продemonстрированы все основные умения. Решены все основные задачи. Выполнены все задания, в полном объеме, но некоторые с недочетами.	Продemonстрированы все основные умения, решены все основные задачи с отдельными несущественными недочетами, выполнены все задания в полном объеме.	Продemonстрированы все основные умения, решены все основные задачи. Выполнены все задания, в полном объеме без недочетов
<u>Навыки</u>	Отсутствие владения материалом. Невозможность оценить наличие навыков вследствие отказа обучающегося от ответа	При решении стандартных задач не продемонстрированы базовые навыки.  Имели место грубые ошибки.	Имеется минимальный набор навыков для решения стандартных задач с некоторыми недочетами.	Продemonстрированы базовые навыки при решении стандартных задач с некоторыми недочетами	Продemonстрированы базовые навыки при решении стандартных задач без ошибок и недочетов.	Продemonстрированы навыки при решении нестандартных задач без ошибок и недочетов.	Продemonстрирован творческий подход к решению нестандартных задач.

### Шкала оценки при промежуточной аттестации

Оценка	Уровень подготовки
Зачтено	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «превосходно»
	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»
	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»
	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»
	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	Хотя бы одна компетенция сформирована на уровне «неудовлетворительно», ни одна из компетенций не сформирована на уровне «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

Вопросы	Код формируемой компетенции
<i>Основные понятия и свойства. Разведочный анализ</i>	
1. Матрица данных, типы признаков (числовой, ординальный, номинальный), срезы данных (продольный, поперечный). Описательная статистика числового признака (среднее, дисперсия, стандарт). Центрирование и стандартизация данных. Мода, медиана, гистограмма, квартили, выброс.	ПК-6
2. Ковариация и корреляция. Понятия строгой, сильной, средней, слабой положительной и отрицательной корреляции и ее отсутствия. <u>Свойства коэффициента ковариации и свойства коэффициента корреляции Пирсона для исходных, центрированных, стандартизованных признаков</u> . Анализ диаграмм рассеивания. Ложная корреляция, выбросы, корреляция и нелинейная зависимость.	ПК-6
3. Статистика признаков. Вектор средних значений. Матрица ковариаций. Матрица корреляций. Матрица ковариаций двух наборов признаков (определения). Свойства матрицы ковариаций для центрированных и стандартизованных признаков. Способы записи и вычисления матрицы ковариаций и корреляций.	ПК-6
4. Статистика случайных векторов. Математическое ожидание случайного вектора. Матрицы ковариаций и корреляций случайного вектора. Матрица ковариаций двух случайных векторов (определения). Запись матрицы ковариаций и корреляций через математические ожидания. Свойства матрицы ковариаций стандартизованного случайного вектора.	ПК-6

<i>Вопросы</i>	<i>Код формируемой компетенции</i>
5. Теорема о дисперсии линейной комбинации независимых случайных величин. Теорема о линейных свойствах ковариации случайных величин. Теорема о дисперсии линейной комбинации двух признаков. Теорема о линейных свойствах ковариации признаков (формулировки). Ортогональные матрицы и их свойства (формулировки).	ПК-6
<i>Кластерный анализ</i>	
6. Понятие суммарной дисперсии признаков (для матрицы данных). <u>Геометрические свойства суммарной дисперсии.</u>	ПК-6
7. Задача кластерного анализа. Таблица принадлежности объектов кластерам. Понятия центра кластера, расстояния от объекта до кластера, расстояния между кластерами, расстояния от объекта до «своего кластера». <i>Понятие показателя качества кластерного решения D.</i>	ПК-6
8. <u>Геометрические свойства показателя качества кластерного решения D.</u>	ПК-6
9. <u>Правило разложения дисперсии на внутригрупповую и междугрупповую. Статистический смысл показателя D и его свойства.</u>	ПК-6
10. Алгоритм и основное свойство метода K-средних. Принцип построения цепочки кластерных решений и <i>выбора решения.</i>	ПК-6
11. Метод иерархического кластерного анализа. Способы определения сходства объектов, сходства объекта и группы объектов, сходство двух групп объектов (способы выбора расстояния, способы выбора правила связи).	ПК-6
12. Алгоритм метода иерархического кластерного анализа. Схема объединения, расстояние объединения, функция объединения. Критерий выбора кластерного решения.	ПК-6
13. Сравнение методов K-средних и иерархического метода кластеризации.	ПК-6
<i>Метод главных компонент</i>	
14. Условия, цели и схема применения метода главных компонент (МГК). Критерии качества снижения размерности (доля объясненной суммарной дисперсии, общности, <i>по лекциям</i> ). Матрица весовых коэффициентов главных компонент (ВКГК), смысл ее коэффициентов. Выбор названий главных компонент.	ПК-6
15. Исходные данные для МГК, свойства матрицы корреляций (R), используемые при построении метода, <u>связь матриц R и Z</u> , определение главных компонент. Определение матрицы значений главных компонент. <u>Размерность, свойства и связь матриц S, Z, W.</u>	ПК-6
16. Свойства главных компонент: <u>утверждения о средних и дисперсиях ГК.</u>	ПК-6
17. Суммарная дисперсия набора признаков и ее геометрический смысл. <u>Утверждение о суммарной дисперсии ГК</u> , его геометрический смысл.	ПК-6
18. <u>Утверждение об отсутствии корреляции главных компонент.</u>	ПК-6
19. <u>Утверждения о свойствах матрицы ВКГК</u> (весовых коэффициентов главных компонент).	ПК-6
20. Определение общностей через коэффициенты матрицы ВКГК, <u>их смысл и свойства</u> (в т.ч. связь с определением из лекций).	ПК-6
21. Определение показателя $\delta$ через коэффициенты матрицы ВКГК, <u>его смысл и свойства</u> (в т.ч. связь с определением из лекций).	ПК-6

Вопросы	Код формируемой компетенции
22. Выбор количества главных компонент при решении задач снижения размерности.	ПК-6
<i>Факторный анализ</i>	
23. Модель факторного анализа в канонической и векторной записи. <u>Основное свойство факторной модели</u> . Воспроизведенная матрица ковариаций.	ПК-6
24. Общности, характерности и доля объясненной дисперсии. <u>Теорема о ковариациях исходных признаков и общих факторов</u> .	ПК-6
25. Ортогональные матрицы. <u>Теорема о вращения факторов</u> . Инварианты при вращении. Цели вращения факторов.	ПК-6
26. Факторный анализ: <u>теорема о линейном преобразовании признаков</u> . Модели для стандартизованных признаков.	ПК-6
27. Факторная модель для стандартизованных признаков в канонической и векторной записи. <u>Основное свойство и его обоснование</u> . Воспроизведенная матрица корреляций.	ПК-6
28. <u>Общности, характерности и доля объясненной дисперсии (определения, свойства и их доказательства)</u> .	ПК-6
29. <u>Теорема о корреляции признаков и общих факторов</u> . Диаграммы факторных нагрузок. Принципы вращения факторов.	ПК-6
30. Модель для стандартизованных признаков. Задачи и схема исследования. Особенности применения модели. <i>Критерии качества решения</i> .	ПК-6
31. Геометрический смысл корреляции признаков в пространстве объектов и пространстве признаков. <i>Геометрический смысл факторного анализа</i> .	ПК-6
32. Проблемы поиска факторных нагрузок. <u>Теорема о «диагональных» условиях нормировки</u> .	ПК-6
33. <u>Принцип главных факторов (теорема)</u> . Формулировки теорем о ранге матричного произведения и спектральном разложении.	ПК-6
34. Обоснование ограничений на число общих факторов.	ПК-6
35. <u>Оценка факторных нагрузок методом главных факторов</u> : постановка задачи, алгоритм и критерии останова. Оценка числа общих факторов и начальная оценка общностей.	ПК-6
36. <u>Оценка факторных значений</u> : матрица данных как запись реализаций случайных векторов, задача на отыскание факторных значений, регрессионный способ оценки. <u>Свойства матриц</u> $\hat{F}$ , $V(\hat{F})$ и $\hat{A}$ . Факторные значения и вращение факторов.	ПК-6
<i>Представление математических моделей в табличном виде</i>	
37. Многомерная статистическая модель «Территория. Год. Показатели». Анализ структуры, динамики, территориальной динамики. Средства визуализации анализа.	ПК-6
38. Модели поколения и населения. Показатели таблиц дожития.	ПК-6
39. Статистический анализ таблиц дожития.	ПК-6



### 5.2.2. Типовые тестовые задания для оценки сформированности компетенции ПК-6, индикатор ПК-6.1

#### Образцы тестовых вопросов:

##### *Вопрос 1*

**Тип вопроса:** ввод значения

Формулировка вопроса: Файл данных *Практикум\_1.sav*. Формат выдачи: SPSS. (формат «целое число»)

**Ответ 1:** число исследуемых признаков

Верное значение: 4

**Ответ 1:** число объектов

Верное значение: 108

**Ответ 3:** число объектов с пропусками данных

Верное значение: 0

**Ответ 4:** строим кластеры в пространстве размерности

Верное значение: 4

##### *Вопрос 2*

**Тип вопроса:** одиночный выбор

Формулировка вопроса: Файл данных *Практикум\_1.sav*. Формат выдачи: SPSS.

Кластеризация будет проведена

**Варианты ответов:**

- для исходных признаков
- для исходных стандартизованных признаков (+)

##### *Вопрос 3*

**Тип вопроса:** ввод значения

Формулировка вопроса: Файл данных *Практикум\_1.sav*. Формат выдачи: SPSS.

Общий центр в исходных координатах (формат *\*\*.\**)

**Ответ 1:** LE\_m

Верное значение: 62.7

**Ответ 2:** LE\_f

Верное значение: 67.8

**Ответ 3:** KMC

Верное значение: 46.9

**Ответ 4:** KCP

Верное значение: 3.7

##### *Вопрос 4*

**Тип вопроса:** ввод значения

Формулировка вопроса: Файл данных *Практикум\_1.sav*. Формат выдачи: SPSS.

Общий центр в стандартизованных координатах (формат *\*\*.\**)

**Ответ 1:** LE\_m

Верное значение: 0.0

**Ответ 2:** LE\_f

Верное значение: 0.0

**Ответ 3:** KMC

Верное значение: 0.0

**Ответ 4:** KCP

Верное значение: 0.0

##### *Вопрос 5*

**Тип вопроса:** ввод значения

Формулировка вопроса: Файл данных *Практикум\_1.sav*. Формат выдачи: SPSS. (формат *\*.\*\*\**)

**Ответ:** исследуемая суммарная дисперсия

Верное значение: 4.000

### **Вопрос 6**

**Тип вопроса:** ввод значения

Формулировка вопроса:

Файл данных *Практикум\_1.sav*. Формат выдачи: SPSS.

Показатели качества решений (в процентах) (формат \*\*\*.\*)

**Ответ 1:** 2 кластера

Верное значение: 79.4%

**Ответ 2:** 3 кластера

Верное значение: 88.0%

**Ответ 3:** 4 кластера

Верное значение: 91.1%

### **Вопрос 7**

**Тип вопроса:** одиночный выбор

Формулировка вопроса:

Файл данных *Практикум\_1.sav*. Формат выдачи: SPSS.

Для решения на 3 кластера при значимости 0.01 кластерные центры

**Варианты ответов:**

- незначимо отличаются по всем координатам
- по некоторым координатам значимо не отличаются
- по каждой координате отличаются значимо (+)
- нет значимых отличий по всем координатам

### **Вопрос 8**

**Тип вопроса:** ввод значения

Формулировка вопроса: Файл данных *Практикум\_1.sav*. Формат выдачи: SPSS.

Для решения на 3 кластера: стандартизованные координаты кластерного центра самого многочисленного кластера (формат \*.\*\*\*)

**Ответ 1:** LE\_m

Верное значение: 0.762

**Ответ 2:** LE\_f

Верное значение: 0.812

**Ответ 3:** KMC

Верное значение: -0.808

**Ответ 4:** KCP

Верное значение: -0.901

### **Вопрос 9**

**Тип вопроса:** классификация

Формулировка вопроса: Файл данных *Практикум\_1.sav*. Формат выдачи: SPSS.

Для решения на 3 кластера: координата кластерного центра самого многочисленного кластера отличается от координаты центра совокупности

**Ответ 1:** LE\_m, *Классифицирующий термин* – с превосходством

**Ответ 2:** LE\_f, *Классифицирующий термин* – с превосходством

**Ответ 3:** KMC, *Классифицирующий термин* – с недостатком

**Ответ 4:** KCP, *Классифицирующий термин* – с недостатком

### 5.2.3. Типовые задания/задачи для оценки сформированности компетенции ПК-6

#### ❖ Типовые практические задания для оценивания сформированности компетенции ПК-6, индикатор ПК-6.1

Сформулируйте определения и докажите утверждения из следующих разделов дисциплины:

##### *Основные понятия и свойства. Разведочный анализ*

1. Свойства коэффициента ковариации и свойства коэффициента корреляции Пирсона для исходных, центрированных, стандартизованных.

##### *Кластерный анализ*

2. Геометрические свойства суммарной дисперсии.
3. Геометрические свойства показателя качества кластерного решения D.
4. Правило разложения дисперсии на внутригрупповую и междугрупповую.

##### *Метод главных компонент*

5. Утверждения о средних и дисперсиях ГК, о суммарной дисперсии ГК, об отсутствии корреляции ГК, свойствах весовых коэффициентов главных компонент.
6. Определение общностей, их смысл и свойства, определение доли суммарной объясненной дисперсии. Выбор количества главных компонент.

##### *Факторный анализ*

7. Факторная модель для стандартизованных признаков в канонической и векторной записи. Воспроизведенная матрица корреляций. Общности, характеристики и доля объясненной дисперсии. Теорема о корреляции признаков и общих факторов.
8. Теорема о «диагональных» условиях нормировки. Принцип главных факторов.
9. Оценка факторных нагрузок методом главных факторов.
10. Оценка факторных значений.

#### ❖ Типовые практические задания для оценивания сформированности компетенции ПК-6, индикатор ПК-6.2

Стартовый практикум. Первичный анализ данных, анализ структуры и динамики (в EXCEL)

Практикум 1. Первичный анализ данных. Метод К-средних (в SPSS®)

Практикум 2. Метод главных компонент (в SPSS®)

Практикум 3. Факторный анализ (в SPSS®)

Дополнительный практикум. Применение метода главных компонент для анализа данных, упорядоченных по времени (в SPSS®)

Содержание и порядок выполнения заданий, а также комплекты данных, приведены в [1], [3].

##### *Контрольные вопросы к практическим заданиям*

1. Опишите исходные данные.
2. Сформулируйте цель и задачу исследования.
3. Проведите первичный анализ данных.
4. Аргументируйте применение предложенного в задании метода.
5. Опишите понятийный аппарат проведенного исследования, сформулируйте его основные свойства. Комментируйте компьютерную диагностику построения решения.

6. Проверьте корректность применения метода (если нужно, используйте статистический вывод), и определите качество решения.
7. Сформулируйте результаты исследования.

❖ **Типовое практическое проектное задание для оценивания сформированности компетенции ПК-6, индикатор ПК-6.3**

Практикум 4. Самостоятельный анализ данных

Используйте данные в соответствии с *индивидуальным заданием*, полученным от преподавателя.

Поставьте задачу, то есть *определитесь с массивом исследуемых факторов*. Определите цель исследования.

Используйте методы *первичного анализа* и *метод (методы) снижения размерности*.

Метод (кластерный анализ, метод главных компонент, факторный анализ) либо указан в индивидуальном задании, либо нужно выбрать тот, который даст наиболее убедительный (по качеству) результат.

Инструмент анализа: *статистический пакет*.

По проведенной работе подготовьте *отчет*.

*Отчет начинается* с постановки задачи, описания исследуемых переменных (факторов), указания объема выборки, полноты данных, используемой группировки и т.п. Указывается цель исследования.

*Отчет включает* ключевые показатели снижения размерности, графики и диаграммы, результаты проверки значимости, координаты кластерных центров и т.п. в соответствии с выходными формами компьютерной обработки, изученными в рамках практикума.

*Отчет завершается* описанием (интерпретацией) полученных результатов.

Готовый отчет должен быть дополнен (на полях) *рукописным указанием обозначений и свойств всех существенных характеристик*, используемых в выходных формах компьютерной обработки.

*При подготовке формульных фрагментов отчета рукописная форма обязательна*. При подготовке постановочной и заключительной части *рукописная форма приветствуется*.

При отсутствии доступа к программным средствам анализируется материал с описанием файла данных и выходными формами, указанными в *индивидуальном задании*.

Отчет высылают в формате сканированного документа или в формате doc, pdf.

Примерные комплекты данных приведены в [1].

*Контрольные вопросы к практическому заданию (проекту)*

1. Опишите исходные данные, целесообразность их фильтрации и (или) группировки. Опишите первоначальные и окончательные цели и задачи исследования (их может быть несколько).
2. Сформулируйте результаты первичного анализа данных.
3. Аргументируйте выбор методов для дальнейшего исследования. Опишите их понятийный аппарат.
4. Опираясь на компьютерную диагностику, опишите этапы применения методов.
5. Проверьте корректность применения методов (если нужно, используйте статистический вывод), определите качество предложенных решений.
6. Структурируйте полученные результаты, сформулируйте выводы.
7. Проведите сравнительный анализ использованных подходов и методов.

#### 5.2.4. Типовые вопросы для собеседования по проверке компетенции ПК-6

##### *Собеседование по результатам выполнения практических заданий – Практикум*

1. Опишите исходные данные.
2. Сформулируйте цель и задачу исследования.
3. Проведите первичный анализ данных.
4. Аргументируйте применение предложенного в задании метода.
5. Опишите понятийный аппарат проведенного исследования, сформулируйте его основные свойства. Комментируйте компьютерную диагностику построения решения.
6. Проверьте корректность применения метода (если нужно, используйте статистический вывод), и определите качество решения.
7. Сформулируйте результаты исследования.

##### *Собеседование по результатам выполнения проекта (в рамках Практикума)*

1. Опишите исходные данные, целесообразность их фильтрации и (или) группировки. Опишите первоначальные и окончательные цели и задачи исследования (их может быть несколько).
2. Сформулируйте результаты первичного анализа данных.
3. Аргументируйте выбор методов для дальнейшего исследования. Опишите их понятийный аппарат.
4. Опираясь на компьютерную диагностику, опишите этапы применения методов.
5. Проверьте корректность применения методов (если нужно, используйте статистический вывод), определите качество предложенных решений.
6. Структурируйте полученные результаты, сформулируйте выводы.
7. Проведите сравнительный анализ использованных подходов и методов.

#### 5.2.5. Типовые задания (оценочные средства), выносимые на зачет

##### *На зачет по дисциплине «Методы моделирования и прикладной анализ данных – 2» выносятся:*

1. Собеседование по результатам выполнения практических заданий теоретического характера, см. п. 5.2.4.
2. Тестирование, см. п. 5.2.2.
3. Собеседование по результатам выполнения проекта, см. п. 5.2.4.
4. Собеседование по вопросам, указанным в п. 5.1.

#### 6. Учебно-методическое и информационное обеспечение дисциплины

##### а) основная литература:

1. Стронгина Н.Р. Методы моделирования: анализ данных. ЭУК, учебно-методический комплекс. Фонд электронных образовательных ресурсов ННГУ. Н. Новгород, 2014. Идентификационный номер 814Е.14.08 <https://e-learning.unn.ru/enrol/index.php?id=827>. – Требуется регистрация.
2. Российский статистический ежегодник. Росстат (ежегодное официальное издание). [http://www.gks.ru/bgd/regl/b16\\_13/Main.htm](http://www.gks.ru/bgd/regl/b16_13/Main.htm), сайт Федеральной службы государственной статистики.

##### б) дополнительная литература:

3. Стронгина Н.Р. Методы моделирования: регрессионный анализ. ЭУК, учебно-методический комплекс. Фонд электронных образовательных ресурсов ННГУ. Н. Новгород, 2014. Идентификационный номер 793Е.14.08 <https://e-learning.unn.ru/enrol/index.php?id=827>. – Требуется регистрация.
4. Стронгина Н.Р., Марчева И.А. - Регрессионный анализ в экономических приложениях: учеб. пособие. – Н. Новгород: Изд-во Нижегород. гос. ун-та, 2003. – 214 с. (5 экз.)
5. Романов В.П. Интеллектуальные информационные системы в экономике: Учебное пособие. /Под редакцией д.э.н., профессора Н.П. Тихомирова. – М.: Изд-во «Экзамен», 2003. – 496 с. (108 экз).

в) программное обеспечение и Интернет-ресурсы (в соответствии с содержанием дисциплины):

Специализированное программное обеспечение SPSS или Statistica (лицензия), а также MS Excel, Word.

## **7. Материально-техническое обеспечение дисциплины**

Помещения представляют собой учебные аудитории для проведения учебных занятий, предусмотренных программой, оснащенные оборудованием и техническими средствами обучения: терминал-класс с лицензионным программным обеспечением (специализированное лицензионное программное обеспечение SPSS или Statistica, а также Excel, Word); применяются презентации и проекционное оборудование (персональный компьютер, проектор, экран).

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

Программа составлена в соответствии с требованиями ОС ННГУ 01.03.02 Прикладная математика и информатика.

Автор: к.ф.-м.н., доцент кафедры ДУМиЧА Стронгина Н.Р.

Рецензент: д.т.н., профессор НГТУ им. Р.Е. Алексеева Ломакина Л.С.

Заведующий кафедрой ДУМиЧА: д.ф.-м.н. Калинин А.В.

Программа одобрена на заседании методической комиссии института информационных технологий, математики и механики

от 30 ноября 2022 года, протокол № 3.