

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ

РОССИЙСКОЙ ФЕДЕРАЦИИ

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

Факультет социальных наук

УТВЕРЖДЕНО

решением Ученого совета ННГУ

протокол № 1 от 16.01.2024 г.

Рабочая программа дисциплины

**Использование программ
анализа данных в экономико-социологических исследованиях**

Уровень высшего образования

Подготовка научных и научно-педагогических кадров

Программа аспирантуры
Экономическая социология

Научная специальность
5.4.2. Экономическая социология

Квалификация (степень)
Исследователь. Преподаватель-исследователь

Форма обучения
Очная, заочная

Нижний Новгород
2024

1. Место дисциплины в структуре ОПОП

Дисциплина 2.1.02.05 Использование программ анализа данных в экономико-социологических исследованиях относится к части, формируемой участниками образовательных отношений образовательной программы.

Целями освоения дисциплины являются освоение аспирантами понятийного аппарата современно компьютерного анализа социологической информации, овладение приемами и методами анализа данных социологических исследований.

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

Знать:

- теоретические основы методов обработки данных (количественных, качественных, комплексного анализа)

Уметь:

- применять современные методы обработки данных (статистические пакеты, программные комплексы дискурс- и контент-анализа)

Владеть:

- навыком отбора и применения необходимых методов обработки данных в социологических исследованиях.

2. Структура и содержание дисциплины

2.1 Трудоемкость дисциплины

	очная
Общая трудоемкость, з.е.	2
Часов по учебному плану	72
в том числе	
аудиторные занятия (контактная работа):	
- занятия лекционного типа	18
- занятия семинарского типа (практические занятия / лабораторные работы)	18
- КСР	2
самостоятельная работа	28
Промежуточная аттестация	зачет

2.2. Содержание дисциплины

(структурированное по темам (разделам) с указанием отведенного на них количества академических часов и виды учебных занятий)

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 (практические занятия/лабора	Всего	

			торные работы), часы		
	о ф о	о ф о	о ф о	о ф о	о ф о
Основные понятия количественного анализа социологических данных Объект, переменная, матрица данных, типы переменных. Преобразования переменных. Типы шкал.	6	2	2	4	2
Оценка качества данных в социологических исследованиях Контроль качества массива данных: Проверка заполняемости. Выборочный контроль кодирования. Проверка на непротиворечивость. Выявление дублей, фальсифицированных анкет и дописанных самостоятельно ответов. Анализ соответствия структуры выборки генеральной совокупности. Анализ качества шкал. Анализ качества структуры инструментария в целом	6	2	2	4	2
Базовые понятия описательной статистики Процентные показатели. Меры центральной тенденции (среднее, мода, медиана). Меры вариации (размах, дисперсия, стандартное отклонение, коэффициент вариации). Линейные и условные распределения.	6	2	2	4	2
Взаимосвязь признаков Коэффициенты связи: Хи-квадрат, t-критерия Стьюдента, коэффициенты Юла, Крамера, Чупрова, Гудмена-Крускала, Спирмена, Кэндалла, Пирсона.	6	2	2	4	2
Понятие ошибки выборки Основные понятия и типы выборочного метода в социологическом исследовании (вероятностные: случайная, механическая, стратифицированная, серийная - и невероятностные выборки: квотная, снежный ком, стихийная, типичных случаев). Оценка репрезентативности выборки в конкретном проекте. Проверка соответствия итоговой выборки исходной модели по структуре и объемам. Коррекция выборки в соответствии с исходной моделью (ремонт выборки).	6	2	2	4	2
Цели, задачи, гипотезы как основа формирования статистического задания к исследованию Формулирование целей, задачи и гипотез исследования. Определение задач, доступных для решения опросными методами. Подбор типов шкал, наиболее оптимальных для каждой задачи.	5	1	2	3	2
Влияние корректности статистической обработки и визуализации данных на формулирование выводов исследования Формулирование идеи как основа для выбора типа используемой статистики. Типичные ошибки при расчетах и их влияние на выводы. Определение целевой аудитории отчета и формата представления данных. Основные принципы визуализации данных. Типичные ошибки визуализации данных и влияние на выводы.	4	1	1	2	2
Принципы построения рейтингов и индексов. Зачем нужны рейтинги и индексы. Типы параметров, которые возможно включить в индекс. Нормирование параметров для включения в индекс. Весовые коэффициенты. .	5	2	1	3	2
Кластерный анализ Виды кластерного анализа. Принцип действия. Основные понятия. Меры расстояния. Правила объединения. Алгоритмы построения. Условия применимости. Способы определения оптимального количества кластеров. Основы интерпретации	6	1	1	2	4
Факторный анализ История появления. Сферы применения. Принцип действия. Основные понятия. Условия применимости. Способы определения оптимального количества факторов. Основы интерпретации	6	1	1	2	4
Регрессионный анализ Основные понятия. Цели регрессионного анализа. Виды регрессии. Определение оптимальной модели. Метод наименьших квадратов. Оценка качества линейной регрессии. Инструменты для построения регрессионного анализа.	6	1	1	2	4
Итого	72	18	18	36	28

3. Учебно-методическое обеспечение самостоятельной работы обучающихся

Для обеспечения самостоятельной работы обучающихся используются:

- электронный курс "Использование программ анализа данных в экономико-социологических исследованиях" (<https://elearning.unn.ru/course/view.php?id=7287>).
- открытый онлайн-курс МООС "нет" (нет).

4. Фонд оценочных средств для текущего контроля успеваемости и промежуточной аттестации по дисциплине (модулю)

4.1 Типовые задания, необходимые для оценки результатов обучения при проведении текущего контроля успеваемости с указанием критериев их оценивания:

4.1.1 Типовые задания

1:

Лекционные занятия: Рассмотрение предложенного ряда вопросов и определение типа шкал для каждого из них: номинальная, порядковая, количественная шкала. Обсуждение возможных вариантов математических расчетов для отображения результатов по каждому исследованию.

Лабораторные занятия и самостоятельная работа обучающихся:

1. Создание структуры матрицы данных в Excel согласно выбранным анкетам и специфике типов переменных в них. Контроль за соблюдением стандартов организации матрицы.
2. Импорт файла с базой данных из Excel в программный пакет анализа данных. Организация подписей меток переменных и свойств данных

2:

Лабораторные занятия и самостоятельная работа обучающихся: на основе ранее подготовленных массивов данных необходимо произвести:

1. Проверку заполняемости и контроль кодирования:
2. Логическую проверку данных (непротиворечивость ответов)
3. Сформировать перечень претензий к массиву данных
4. Предложить возможные пути устранения выявленных нарушений

3:

Лекционные занятия: разбор конкретных кейсов с результатами исследований по вопросам:

1. Какой статистический показатель целесообразно использовать для отражения результатов (частота, %), что может помочь определиться с выводами.
2. От какого объема выборки (целой выборки или подвыборки) целесообразно считать % и почему. Какие выводы позволит сделать тот или иной подход.
3. Какой статистический показатель целесообразно использовать для отражения результатов при анализе центральной тенденции (среднее, мода, медиана и пр.) и почему. Как могут меняться выводы в зависимости от выбранного типа статистики.
4. Оценить, что можно считать «выбросами», требующими исключения из общей статистики. Обосновать, почему необходимо исключение.
5. Рассчитать среднее, дисперсию, стандартное отклонение, коэффициент вариации, для предложенных рядов данных.

Лабораторные занятия и самостоятельная работа обучающихся: на основе ранее подготовленных массивов данных необходимо:

1. Построить линейные распределения по всем параметрам массива данных. Обязательная итоговая статистика: частота, %, меры средней тенденции.
2. Проанализировать сделанное:
 1. Удобны ли полученные таблицы для последующего использования
 2. Корректно ли посчитана статистика, есть ли «выбросы», насколько они влияют на итоговый результат.
3. Создание наборов множественных ответов и построение статистик с их использованием.

4:

Лекционные занятия: разбор конкретных кейсов:

1. неавтоматизированный расчет статистик: Хи-квадрат,
2. оценка значимости результатов предложенных готовых показателей

Лабораторные занятия и самостоятельная работа обучающихся: на основе ранее подготовленных массивов данных необходимо выполнить в программном пакете анализа данных:

1. Расчет условных распределений
2. Оценка статистической значимости наблюдаемых различий с использованием тестов (Хи-квадрат, t-test, ...)

5:

Лабораторные занятия и самостоятельная работа обучающихся: на основе ранее выполненных расчетов необходимо выполнить:

1. расчет ошибки выборки на основе онлайн-калькуляторов
2. оценку соответствия результатов исследования условному квотному заданию
3. расчет поправочных весовых коэффициентов для корректировки выборки.

6:

Лекционные занятия:

Практическая письменная работа 1 – по готовой предложенной анкете...

1. Сформулировать цели, которые мог ставить перед собой исследователь, реализуя исследование именно по такому инструментарию.
2. Сформулировать предположительные гипотезы исследования
3. Сформулировать задачи исследования. Восстановить логику – решение какой задачи какими вопросами решается. Создать таблицу соответствия – какую задачу решает каждый из вопросов.
4. По каждому вопросу прописать виды стат. показателей, необходимых для формирования выводов.

Практическая письменная работа 2 – по готовому техническому заданию

1. Определить, какие из задач исследования возможно решить опросными методами.
2. Придумать вопросы к каждой из задач, решаемых опросными методами. Обязательно применение одновариантных вопросов с номинальной шкалой, поливариантных

По итогам каждой из работ проводится обсуждение типичных ошибок.

71:

Лабораторные занятия и самостоятельная работа обучающихся: на основе выбранного массива данных необходимо:

1. Выбрать группу переменных, которые могут являться ключевыми индикаторами для сегментации целевой аудитории, а также имеют одинаковую размерность шкалы.
2. Провести кластерный анализ иерархическим методом, используя пакет программный пакет анализа данных.
3. Построить условные распределения с оптимальным вариантом кластеризации и описательными характеристиками сегментов (пол, возраст, уровень доходов, ответы на другие вопросы, другие выделенные сегменты и пр.).
4. Описать результаты каждого этапа в формате слайдов презентации: логику построения модели и смысловая интерпретация полученных результатов.

8:

Лабораторные занятия и самостоятельная работа обучающихся: на основе выбранного массива данных необходимо:

1. Выбрать группу переменных, которые могут являться ключевыми индикаторами для сегментации целевой аудитории, а также имеют одинаковую размерность шкалы.
2. Провести кластерный анализ методом К-средних, используя программный пакет анализа данных.
3. Построить условные распределения с оптимальным вариантом кластеризации и описательными характеристиками сегментов (пол, возраст, уровень доходов, ответы на другие вопросы, другие выделенные сегменты и пр.).
4. Описать результаты каждого этапа в формате слайдов презентации: логику построения модели и смысловая интерпретация полученных результатов.

9:

Лабораторные занятия и самостоятельная работа обучающихся: на основе выбранного массива данных необходимо:

1. Выбрать группу переменных, которые могут являться ключевыми индикаторами для сегментации целевой аудитории, а также имеют одинаковую размерность шкалы.
2. Провести кластерный анализ методом К-средних, используя пакет программный пакет анализа данных.
3. Построить условные распределения с оптимальным вариантом кластеризации и описательными характеристиками сегментов (пол, возраст, уровень доходов, ответы на другие вопросы, другие выделенные сегменты и пр.).
4. Описать результаты каждого этапа в формате слайдов презентации: логику построения модели и смысловая интерпретация полученных результатов.

10:

Лабораторные занятия и самостоятельная работа обучающихся: на основе выбранного массива данных необходимо:

1. Выбрать переменные и построить факторную модель
2. Присвоить названия полученным факторам и описать их

3. На основе полученных факторов построить кластерную модель методом К-средних. Присвоить названия кластерам и описать их.
4. Построить условные распределения с оптимальным вариантом кластеризации и описательными характеристиками сегментов
5. Описать результаты каждого этапа в формате слайдов презентации

11:

Лабораторные занятия и самостоятельная работа обучающихся: на основе доступных статистических показателей ресурса www.fedstat.ru необходимо:

1. Выбрать **показатель, развитие которого необходимо спрогнозировать**
2. Сформулировать гипотезу, какие другие показатели могут в высокой степени коррелировать с выбранным зависимым показателем
3. Через последовательную оценку нескольких показателей (с помощью метода парной регрессии, встроенного в функцию построения точечных графиков в пакете MS Excel) найти предиктор, наилучшим образом определяющий развитие зависимой переменной.
4. Описать полученный результат (какая модель оказалась наиболее оптимальной, насколько сильна зависимость, каков характер зависимости и как это можно проинтерпретировать)

Лабораторные занятия и самостоятельная работа обучающихся: На основе предложенного массива данных:

1. Выбрать зависимую и набор независимых переменных.
2. Провести регрессионный анализ методом множественной регрессии в программном пакете анализа данных.
3. Описать полученный результат – оценить качество модели и дать ее смысловую интерпретацию.

Тестовые задания для текущего контроля

Тестовые задания:

1. Установите соответствие:

1. Программа наблюдения	a совокупность социально-экономических явлений и процессов, которые подлежат исследованию, или точные границы, в пределах которых будут регистрироваться статистические сведения
2. Критический момент наблюдения	b период, в течение которого осуществляется сбор данных, заполнение бланков программы наблюдения
3. Период наблюдения	c время, в течение которого проводится обследование по разработанной программе
4. Время наблюдения	d совокупность действий, необходимых для достижения цели исследования
5. Объект наблюдения	e постановка научной проблемы, определение свойств и тенденций общественного явления, подлежащих анализу
6. Цель исследования	f момент или отрезок времени, по состоянию на который проводится регистрация значений признаков по каждой единице наблюдения. В зависимо-

	сти от целей и задач исследования, особенностей структуры совокупности, предмета исследования критическим моментом может быть дата (день, час), неделя, месяц и т. п.
7. Задачи исследования	g перечень вопросов, по которым собираются сведения, либо перечень признаков и показателей, подлежащих регистрации. Программа наблюдения оформляется в виде бланка (анкеты, формуляра), в который заносятся первичные сведения. Необходимым дополнением к бланку является инструкция (или указания на самих формулярах), разъясняющая смысл вопросов. Состав и содержание вопросов программы наблюдения зависят от задач исследования и от особенностей изучаемого общественного явления

2. Установите соответствие:

1. Номинальная шкала	a имеет фиксированный нуль, а единица измерения в ней может быть произвольной
2. Порядковая (ранговая) шкала	b устанавливает отношения предпочтений между вариантами значений признака (например, группы населения по доходам: нищие, бедные, с доходами ниже среднего уровня, со средними доходами, с доходами выше среднего, богатые). Эта шкала также не имеет нуля (начала отсчета) и единицы измерения
3. Количественная шкала	c устанавливает отношения следования между интервалами значений признака. Имеет произвольный нуль и произвольную единицу измерения
4. Интервальная шкала	d перечень значений атрибутивного признака (например, телефонный справочник). Эта шкала не имеет ни нуля (начала отсчета), ни предпочтений, ни единицы измерения

3. К системе порядковых статистик относятся (отметьте все правильные варианты ответа):

- 1) медиана;
- 2) квартили;
- 3) квинтили;
- 4) децили;
- 5) проценти;
- 6) все выше перечисленное.

4. Какие порядковые статистики делят совокупность на две равные части?

- 1) медиана;
- 2) квартили;
- 3) квинтили;
- 4) децили;
- 5) проценти;
- 6) все выше перечисленное.

5. Выборочная мода (M_o):

- 1) элемент выборки, встречающийся с наибольшей частотой (для унимодального – одновершинного распределения);
- 2) число, которое делит вариационный ряд на две части, содержащие одинаковое количество элементов

6. Установите соответствие между видом группировки и способом ее формирования:

иерархическая	значения второго признака определяются областью значений первого
неиерархическая	не существует строгой зависимости значений второго признака от первого
первичная	составлена на основе первичных данных
вторичная	результат перегруппировки ранее уже сгруппированного материала

7. По заданному значению одной переменной рассчитать наиболее вероятное значение другой переменной позволяет:

- 1) корреляционный анализ
- 2) простой регрессионный анализ
- 3) множественный регрессионный анализ
- 4) факторный анализ
- 5) дисперсионный анализ

8. Предсказать принадлежность объектов к двум или более непересекающимся группам позволяет:

- 1) регрессионный анализ
- 2) кластерный анализ
- 3) дискриминантный анализ
- 4) факторный анализ

9. Предсказать значения зависимой переменной, измеренной в номинальной шкале, позволяет:

- 1) корреляционный анализ
- 2) дискриминантный анализ
- 3) регрессионный анализ
- 4) факторный анализ

10. Предсказать значения зависимой переменной, измеренной в дихотомической шкале, позволяет:

- 1) корреляционный анализ
- 2) простой регрессионный анализ
- 3) логистическая регрессия
- 4) множественная регрессия
- 5) факторный анализ

11. Сжатие информации является главной целью анализа:

- 1) корреляционного
- 2) дисперсионного
- 3) факторного

12. Факторы, влияющие на результат, известны. Выясняется существенность их влияния в анализе:

- 1) корреляционном
- 2) дисперсионном
- 3) факторном

13. В факторном анализе доля общей дисперсии, определяемая данным фактором, называется:

- 1) собственным значением
- 2) нагрузкой
- 3) коэффициентом корреляции

14. При факторном анализе связь между переменной и фактором отражает:

- 1) собственное значение
- 2) нагрузка
- 3) коэффициент корреляции

15. В факторном анализе значение нагрузки переменной лежит в пределах:

- 1) от 0 до $+\infty$
- 2) от -1 до +1
- 3) от $-\infty$ до 0
- 4) от $-\infty$ до $+\infty$
- 5) от -1 до 0
- 6) от 0 до +1

16. Первым этапом процедуры факторного анализа является:

- 1) вращение факторов
- 2) извлечение факторов
- 3): интерпретация факторов
- 4) вычисление корреляционной матрицы для всех переменных

17. Вторым этапом процедуры факторного анализа является:

- 1) вращение факторов
- 2) извлечение факторов
- 3) интерпретация факторов
- 4) вычисление корреляционной матрицы для всех переменных

18. Третьим этапом процедуры факторного анализа является:

- 1) вращение факторов для создания упрощенной структуры
- 2) извлечение факторов
- 3) интерпретация факторов
- 4) вычисление корреляционной матрицы для всех переменных

19. Заключительным этапом процедуры факторного анализа является:

- 1): вращение факторов
- 2) извлечение факторов
- 3) интерпретация факторов
- 4) вычисление корреляционной матрицы для всех переменных

20. Установите соответствие:

методы извлечения факторов	метод главных факторов, метод максимального правдоподобия
методы вращения факторов	метод варимакс (Varimax), метод квартимакс (Equamax)

21. Замена большого числа исходных переменных меньшим числом факторов является целью:

- 1): кластерного анализа
- 2) факторного анализа

22. Для уменьшения числа объектов путем их группировки используется:

- 1) кластерный анализ
- 2) факторный анализ

23. Первым этапом кластерного анализа является:

- 1) выбор способа измерения расстояния между кластерами
- 2) формирование кластеров
- 3) интерпретация результатов
- 4) выбор переменных-критериев для кластеризации

24. Вторым этапом кластерного анализа является:

- 1) выбор способа измерения расстояния между кластерами
- 2) формирование кластеров
- 3) интерпретация результатов
- 4) выбор переменных-критериев для кластеризации

25. Третьим этапом кластерного анализа является:

- 1) выбор способа измерения расстояния между кластерами
- 2) формирование кластеров
- 3) интерпретация результатов
- 4) выбор переменных-критериев для кластеризации

26. Методами формирования кластеров являются:

- 1) метод главных компонент
- 2) метод слияния
- 3) метод дробления
- 4) метод максимального правдоподобия

27. Установите соответствие:

методы кластеризации	внутригрупповое связывание, одиночное связывание, полное связывание
способы измерения расстояния между кластерами	квадрат Евклидова расстояния, косинус, корреляция Пирсона

Критерии оценивания (оценочное средство - Задания)

Оценка	Критерии оценивания
зачтено	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»

4.2. Описание шкал оценивания результатов обучения по дисциплине при промежуточной аттестации

Шкала оценивания сформированности компетенций

Уровень сформированности	плохо	неудовлетворительно	удовлетворительно	хорошо	очень хорошо	отлично	превосходно
--------------------------	-------	---------------------	-------------------	--------	--------------	---------	-------------

компетенций (индикатора достижения компетенций)	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. Ошибок нет.	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 и 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 базовых навыков. 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 оценивания при промежуточной аттестации

Оценка	Уровень подготовки
--------	--------------------

зачтено	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
	отлично	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «отлично».
	очень хорошо	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «очень хорошо»
	хорошо	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «хорошо».
	удовлетворитель	Все компетенции (части компетенций), на формирование которых направлена
	но	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».
	плохо	Хотя бы одна компетенция сформирована на уровне «плохо»

4.3 Типовые контрольные задания или иные материалы, необходимые для оценки результатов обучения на промежуточной аттестации

4.3.1 Типовые задания, выносимые на промежуточную аттестацию:

Зачёт

Критерии оценивания (Задания - Зачёт)

Оценка	Критерии оценивания
зачтено	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»

Задания к зачету:

Файл vooric_Поселенческая выборка.sav. В файле собраны данные из 7 городов, вам *достанется один город (см. приложение ниже)*. Исследование посвящено проблемам отношения населения исторический городов России к культурному наследию и вовлечению граждан в процессы сохранения этого наследия.

По заданному городу:

1. Описать линейное распределение по вопросам: 2,4,5,6,13,17 (текст, таблицы)

2. Выполнить факторный анализ: построить факторную модель по любому блоку вопросов из списка (12,19,20), рассчитать и описать факторные группы с точки зрения вопроса, по которому строилась модель, вопросов 2,4,5,6,13,17 и социально-демографических особенностей группы.

Перед построением факторной модели проанализировать шкалу, ее направленность, состав и т.д.

Необходимые таблицы и графики:

- Таблица полной объясненной дисперсии
- График нормализованного стресса
- Матрица повернутых компонент
- Размеры факторных групп

Таблицы линейки должны быть в тексте работы, таблицы факторных групп – в приложении к работе.

В тексте по факторному анализу должно быть краткое введение (что типологизируем) и выводы («портрет» группы).

1. Выполнить иерархический кластерный анализ переменных по любому блоку вопросов из списка (9,11,18). Определить оптимальное число кластеров, дать им название и краткое описание.

Необходимые таблицы и графики:

- Шаги агломерации
- Дендрограмма с использованием межгрупповых связей
- Таблица принадлежности к кластерам

В тексте по кластерному анализу должно быть краткое введение (ответ на вопрос-классификация чего?). Должны присутствовать название групп и их краткое описание.

1. Курсовая работа должна быть оформлена в соответствии с образцом оформления практических работ

Требования к итоговой письменной работе:

Формат результата: документ Word с диаграммами, таблицами и выводами

Структура документа:

1. Титульный лист
2. Оглавление
3. Методология исследования:
 - а. Описание предполагаемых целей и задач исследования
 - б. Описание характеристик выборочной совокупности
4. где возможно: Портрет опрошенных
5. Основная часть с выделенными смысловыми блоками
6. Выводы и рекомендации
7. Приложения (как минимум – анкета, плюс доп. таблицы - если необходимо)

8. Методологическое резюме по возможностям использованных методов, шкал: по каким вопросам целесообразно изменить размерность шкал, скорректировать формулировки, какие вопросы сделать сразу закрытыми или наоборот открытыми и пр.

Общие требования

1. Логичная структура, осмысленные заголовки и группировка

вопросов 2. Внутри каждого логического блока:

- a. Описание результатов в целом (% и/или средние по всем вопросам, стандартное отклонение, ошибка выборки). Визуализация с помощью диаграмм подходящего типа.
 - b. Условные распределения (выбрать не менее 3 критериев для сравнения, рассчитать условные распределения; оценить статистическую значимость наблюдаемых различий; Оставить только те результаты, для которых выявлена статистическая значимость различий. Обосновать, привлекая данные тестов (Хи-квадрат, t-test).
 - c. Визуализация с помощью диаграмм или таблиц
3. Текстовые выводы к таблицам и диаграммам
4. Обязательно на каждом слайде (или под каждой таблицей или диаграммой, если % считался от разного числа респондентов) должен быть указан объем выборки и максимальная ошибка выборки
5. Аккуратное и грамотное оформление
- a. единая цветовая гамма,
 - b. стандартная для всего отчета размерность и тип шрифтов,
 - c. аккуратное выравнивание и грамотная размерность диаграмм,
 - d. отсутствие грамматических ошибок.
6. Раздел 1. Иерархический кластерный анализ (или методом К-средних – по выбору)
- a. Общая методология исследования
Раздел должен включать в себя описание целевой аудитории, объема выборки, метод исследования (получения информации), цель проведения сегментирования.
 - b. Описание модели
Метод, с помощью которого построена кластерная модель, какие переменные отобраны для построения модели и что они характеризуют, по какому принципу были отобраны именно эти переменные. Кроме того, необходимо привести таблицы с линейным распределением ответов по переменным, включенным в модель. Также необходимо описать, проводились ли какие-либо манипуляции со шкалами или использована исходная шкала. Если проводились, то обосновать, для чего это было сделано. Описать метод формирования кластеров, а также метод определения расстояний между кластерами, обосновать выбор.

Описать алгоритм определения оптимального количества кластеров. Обосновать итоговый результат.
 - c. Описание сегментов
Описать полученную кластерную модель: какое количество групп было выделено, по какому основанию они были выделены, каковы их размеры и чем отличаются друг от друга. Необходимо привести таблицы или линейчатые диаграммы с условным распределением ответов по кластеризующим параметрам в зависимости от принадлежности к тому или иному кластеру. Также необходимо провести анализ сходства или различий выделенных кластеров по прочим характеристикам

(социально-демографическим и другим существенным параметрам исследования). Необходимо привести таблицы(у) (или линейчатые диаграммы), подтверждающие наличие различий.

d. **Резюме (в соответствии с целями проведения сегментирования)**

Какие группы были выявлены, какие являются наиболее многочисленными, какую долю составляет каждая из них, какие рекомендации можно сформировать в соответствии с целями исследования и полученными результатами сегментации.

7. Раздел 2. Факторный анализ.

a. **Общая методология исследования**

Раздел должен включать в себя описание целевой аудитории, объема выборки, метод исследования (получения информации), цель проведения факторного анализа, цель проведения кластерного анализа по итогам факторного.

b. **Описание факторной модели**

Описать, какие переменные были отобраны для построения модели. Необходимо привести таблицу с линейным распределением ответов по переменным, включенным в модель. Также необходимо описать, проводились ли какие-либо манипуляции со шкалами или использована исходная шкала. Если проводились, то обосновать, для чего это было сделано.

Описать, какой метод факторного анализа использован, какое правило применено к пропущенным значениям. Какой метод был использован для определения оптимального количества факторов. Какое решение по количеству факторов было итоговым и почему. Описать, какую объясняющую способность имеет модель. Дать заключение о качестве модели. Обязательно необходимо привести таблицу с собственными значениями факторов и % объясненной дисперсии.

Провести смысловую интерпретацию полученной модели: какие факторы были выявлены, какую смысловую нагрузку несет каждый из них. Обязательно необходимо привести матрицу перевернутых компонент с факторными нагрузками, отсортированными по убыванию абсолютных значений и подписанными названиями факторов. Описать полученные факторы текстом.

c. **Кластерная модель по итогам факторного анализа**

Описать метод, которым была построена кластерная модель. Обосновать итоговое количество кластеров, описать их размер, подкрепив таблицами или диаграммами, описывающими все рассмотренные модели и иллюстрирующими оптимальность выбранного варианта. Описать итоговую модель: присвоить названия кластерам, описать чем они отличаются друг от друга. Обязательно привести таблицу с конечными центрами кластеров. Кластеры должны иметь названия, переменные на основе факторных нагрузок – тоже. В описании групп необходимо использовать выявленные социально-демографические различия, различия по другим параметрам, которые подчеркивают общую идею, суть выделенной группы. Представленные выводы необходимо подкрепить таблицами или диаграммами.

d. **Резюме (в соответствии с обозначенными целями анализа)**

Какие типы (мотивации, поведения и пр.) были выявлены. Какие из них являются наиболее распространенными, у какой доли опрошенных встречаются, какие группы являются носителями тех или иных идей, чем отличаются друг от друга.

Сформулировать рекомендации в соответствии с целями исследования и полученными результатами.

8. Раздел 3. Регрессионный анализ

а. Общая методология исследования

Раздел должен включать в себя обоснование выбора прогнозируемого (зависимого) показателя и набора независимых, цель проведения регрессионного анализа, ссылки на источник информации.

б. Описание рассмотренных моделей

Раздел должен включать в себя таблицы и графики, иллюстрирующие наличие или отсутствие корреляции между рассматриваемыми параметрами. Также необходимо описать оптимальную модель, силу и характер зависимости между показателями, проинтерпретировать результат.

с. Резюме (в соответствии с обозначенными целями анализа)

Оценочное средство - Контрольные вопросы к зачету

1. Основные виды информации. Основные требования к информации в социологических и маркетинговых исследованиях. На каких этапах может теряться информация и почему. Причины непоставимости информации.
2. Что такое объект исследования, матрица данных, переменная. Какие бывают типы переменных и типы шкал.
3. Какие виды базовых математических расчетов можно проводить с номинальной, порядковой и количественной шкалой. Зачем может быть необходимо преобразовывать переменные и каким образом это можно делать.
4. Основные аспекты, по которым проверяется качество матрицы, а также возможные способы ее корректировки.
5. Каким образом можно оценить соответствие выборки структуре генеральной совокупности? Зачем нужен анализ качества шкал и каким образом его можно реализовать?
6. Что такое линейное и условное распределение. В каких случаях применяется тот и другой вариант. От чего можно считать % при расчете условных распределений. Привести конкретный пример и пояснить, как это влияет на выводы.
7. Что такое мода? В каких случаях она предпочтительнее среднего арифметического значения?
8. Что такое медиана? Как она рассчитывается – приведите пример на основе числового ряда? В каких случаях она предпочтительнее среднего арифметического значения?
9. Что такое стандартное отклонение? В каких случаях может быть необходим расчет данного показателя. Приведите примеры.
10. Что такое нормальное распределение? При каких расчетах может быть важно, нормально распределение или нет?
11. Для чего может требоваться перекодировка переменных: изменение направленности шкалы или ее укрупнение? Можно ли укрупнять номинальную шкалу?
12. Что такое Хи-квадрат? Для чего он используется? Как происходит расчет данного коэффициента (что с чем сравнивается)? Как рассчитывается число степеней свободы? Может ли меняться значение

Хи-квадрат с изменением объема выборки, но при сохранении одинакового процентного распределения показателей внутри сравниваемых групп? Если да, то каким образом?

13. Для чего используется t-критерий Стьюдента? Для каких типов переменных и типов распределений применяется? В каких случаях его применение некорректно? Как интерпретировать полученное значение t-критерия Стьюдента и какой еще показатель позволяет сделать вывод о наличии или отсутствии зависимости?

14. Основные типы вероятностных и невероятностных выборок. Понятие репрезентативности исследования. Что такое ошибка выборки, из чего она складывается и как ее можно рассчитать? Что такое доверительный интервал и доверительная вероятность? Для каких типов выборки расчет ошибки выборки применять не корректно? Как в этом случае можно убедиться в репрезентативности результатов исследования?

15. Основные принципы визуализации данных. Типичные ошибки визуализации и их влияние на выводы.

16. Зачем нужны рейтинги и индексы. Типы параметров, которые возможно включить в индекс. Для чего нужно нормирование параметров и введение весовых коэффициентов для индексов, как их можно рассчитать.

17. Понятие классификации. Основные методы классификации: дискриминантный анализ, группировка, кластерный анализ.

18. Иерархический кластерный анализ. Цели использования. Принцип действия. Условия применимости. Способы определения оптимального количества кластеров. Основные принципы интерпретации.

19. Кластерный анализ методом К-средних. Цели использования. Принцип действия. Условия применимости. Способы определения оптимального количества кластеров. Основные принципы интерпретации.

20. Факторный анализ. Цели использования. Принцип действия. Условия применимости. Способы определения оптимального количества факторов. Основные принципы интерпретации.

21. Регрессионный анализ. Цели использования. Виды регрессии. Определение оптимальной модели, метод наименьших квадратов.

22. Оценка качества линейной регрессии. Инструменты для построения регрессионного анализа.

23. Множественный регрессионный анализ. Алгоритм действий. Интерпретация результатов.

5. Учебно-методическое и информационное обеспечение дисциплины (модуля)

Основная литература:

1. Солдаткин Александр Евгеньевич. Использование SPSS в социологических исследованиях : учебно-методическое пособие / А. Е. Солдаткин, А. А. Варызгина, Ю. Ю. Клемянич ; ННГУ им. Н. И. Лобачевского, Факультет социальных наук. - Нижний Новгород : Изд-во ННГУ, 2023. - 21 с. - Текст : электронный., <https://e-lib.unn.ru/MegaPro/UserEntry?Action=FindDocs&ids=850474&idb=0>.
2. Методы обработки данных маркетинговых исследований с применением SPSS : учебно-методическое пособие. - Воронеж : ВГУ, 2018. - 54 с. - Рекомендовано для студентов всех форм обучения бакалавриата факультета международных отношений. Для направления подготовки 38.03.01 – Экономика. - Библиогр.: доступна в карточке книги, на сайте ЭБС Лань. - Книга из

коллекции ВГУ - Экономика и менеджмент., <https://e-lib.unn.ru/MegaPro/UserEntry?Action=FindDocs&ids=754613&idb=0>.

3. Статистическая обработка данных социологического исследования средствами программы SPSS / Красниковский В.Я. - Москва : Прометей, 2021., <https://e-lib.unn.ru/MegaPro/UserEntry?Action=FindDocs&ids=790947&idb=0>.

Дополнительная литература:

1. Прогнозное моделирование в IBM SPSS Statistics и R: Метод деревьев решений / Груздев А.В. - Москва : ДМК-пресс, 2016., <https://e-lib.unn.ru/MegaPro/UserEntry?Action=FindDocs&ids=659481&idb=0>.

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

<https://nafi.ru/academy/prepodavatelyam-spss/>

7. Материально-техническое обеспечение дисциплины (модуля)

Учебные аудитории для проведения учебных занятий, предусмотренных образовательной программой, оснащены мультимедийным оборудованием (проектор, экран), техническими средствами обучения, компьютерами, специализированным оборудованием: проектор
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 ОС ННГУ по направлению 39.04.01 - Социология.

Автор(ы): Солдаткин Александр Евгеньевич, кандидат социологических наук, Варызгина Алла Александровна, кандидат социологических наук.

Заведующий кафедрой: Теодорович Михаил Леонидович, доктор социологических наук.