

КАЗАНСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ
ИНСТИТУТ СОЦИАЛЬНО-ФИЛОСОФСКИХ НАУК И МАССОВЫХ
КОММУНИКАЦИЙ

Кафедра общей и этнической социологии

М.Ю. ЕФЛОВА, К.А. ОЗЕРОВА

**АНАЛИЗ СОЦИОЛОГИЧЕСКОЙ ИНФОРМАЦИИ
В СТАТИСТИЧЕСКИХ ПРОГРАММАХ**

Учебно-методическое пособие

Казань-2021

УДК 311.2
ББК 60.506

*Принято на заседании учебно-методической комиссии ИСФН и МК
Протокол № 4 от 26 февраля*

Ефлова М.Ю., Озерова К.А.

Анализ социологической информации в статистических программах /
М.Ю. Ефлова, К.А. Озерова. – Казань: Казан.ун-т, 2021 – 54 с.

Учебно-методическое пособие по курсу «Анализ социологической информации в статистических программах» предназначено для студентов, обучающихся по направлению подготовки 39.04.01 – «Социология». Составлено в соответствии с требованиями ФГОС ВО к обязательному минимуму содержания и уровня подготовки и учебным планом по направлению 39.04.01 "Социология" и магистерской программе "Социология управления инновационными процессами". Излагается тематическое содержание курса, предлагаются темы семинарских занятий, примеры тестовых и лабораторных работ и рекомендуемая литература.

Подготовлено на кафедре общей и этнической социологии Института социально-философских наук и массовых коммуникаций КФУ.

© Ефлова М.Ю., 2021

© Озерова К.А., 2021

©Казанский университет,2021

ОГЛАВЛЕНИЕ

1.ОРГАНИЗАЦИОННО-МЕТОДИЧЕСКИЕ УКАЗАНИЯ.....	4
2. РАСПРЕДЕЛЕНИЕ КУРСА ПО ТЕМАМ И ВИДАМ РАБОТ	9
3. СОДЕРЖАНИЕ КУРСА	10
4. ПЛАНЫ СЕМИНАРСКИХ ЗАНЯТИЙ.....	23
5. ПРИМЕР ЛАБОРАТОРНОЙ РАБОТЫ	36
6. ВОПРОСЫ К ЗАЧЕТУ	40
7.ТЕСТЫ ДЛЯ ПРОВЕРКИ ОСТАТОЧНЫХ ЗНАНИЙ.....	42
8. РЕКОМЕНДОВАННАЯ ЛИТЕРАТУРА	53
Основная литература.....	53
Дополнительная литература.....	54

1.ОРГАНИЗАЦИОННО-МЕТОДИЧЕСКИЕ УКАЗАНИЯ

1. Цель курса

Цель курса состоит в том, чтобы познакомить обучающихся по магистерской программе "Социология управления инновационными процессами" с основными принципами сбора, обработки и анализа социологической информации при помощи специализированных программных продуктов (статистических пакетов) в рамках прикладных социологических исследований.

2. Задачи курса

В ходе курса студенты будут ознакомлены с ключевыми принципами организации работы с социологической информацией при помощи статистических пакетов программ. В результате у обучающихся должно сформироваться системное представление об алгоритме действий на каждом из этапов по работе с данными в прикладном социологическом исследовании (преимущественно количественном).

Освоение курса должно способствовать:

– выработке способности обоснованно отбирать и использовать современные информационно-коммуникационные технологии для решения профессиональных задач.

– выработке способности самостоятельно формулировать цели, ставить конкретные задачи научных исследований в фундаментальных и прикладных отраслях социологии управления и решать их с помощью современных исследовательских методов с использованием новейшего отечественного и зарубежного опыта и с применением современной аппаратуры, оборудования, информационных технологий.

3. Место дисциплины (модуля) в структуре ОПОП ВО

Данная дисциплина включена в раздел "Б1.В.04 Дисциплины (модули)" основной профессиональной образовательной программы 39.04.01 "Социология

(Социология управления инновационными процессами)" и относится к вариативной части.

Осваивается на 1 курсе в 2 семестре.

В ходе изучения дисциплины магистранты осваивают правила формирования баз социологических данных, приемы обработки и проверки данных. Раскрываются алгоритмы статистического описания, начала анализа (корреляционного, дисперсионного, кластерного).

Обучение происходит на основе актуальных баз социологических данных полученных из открытых источников (данные статистики, результаты социологических опросов). Определенный акцент сделан на способах и формах эффективного представления полученных в ходе анализа результатов при помощи визуализации (таблиц, графиков и диаграмм).

4. Требования к уровню освоения содержания курса «Социальная эксклюзия / инклюзия: стратегии и сценарии социальной политики»

Магистрант окончивший курс «Анализ социологической информации в статистических программах» должен:

Знать релевантные для решения поставленной задачи источники информации, включая национальные и международные базы данных, электронные библиотечные системы, специализированные пакеты прикладных программ. Знать специфику и этапы подготовки методических документов для проведения инструктажа персонала по сбору информации: интервьюеров, кодировщиков, наблюдателей с применением современной аппаратуры, оборудования, информационных технологий.

Уметь производить поиск социологической информации, необходимой для решения поставленной задачи, получать на ее основе социологические данные. Уметь создавать и поддерживать нормативно-методическую и информационную базу исследований по заданной теме. Уметь осуществлять

самостоятельную обработку и анализ социологических данных, с применением стратегий, технологий, ресурсов для проведения научных исследований в фундаментальных и прикладных областях социологии управления с использованием новейшего отечественного и зарубежного опыта

Владеть статистическими процедурами при использовании специализированных пакетов прикладных программ (таких, как MS Excel, Eviews, Stata, SPSS). Владеть современными информационно-коммуникационными технологиями и программными средствами для сбора, обработки и представления информации при решении профессиональных и научно-исследовательских задач.

5. Методические указания для обучающихся по освоению дисциплины

Освоение дисциплины предполагает выполнение обучающимися нескольких видов работ.

Посещение лекции. В ходе лекционных занятий следует вести конспектирование учебного материала. Обращать внимание на категории, формулировки, раскрывающие содержание тех или иных явлений и процессов, научные выводы и практические рекомендации, положительный опыт в ораторском искусстве. Желательно оставить в рабочих конспектах поля, на которых делать пометки из рекомендованной литературы, дополняющие материал прослушанной лекции, а также подчеркивающие особую важность тех или иных теоретических положений. Задавать преподавателю уточняющие вопросы с целью уяснения теоретических положений, разрешения спорных ситуаций.

Перед очередной лекцией необходимо просмотреть по конспекту материал предыдущей лекции. При затруднениях в восприятии материала следует обратиться к основным литературным источникам. Если разобраться в

материале опять не удалось, то обратитесь к лектору (по графику его консультаций) или к преподавателю на практических занятиях.

Практические занятия. В ходе подготовки к семинарам следует изучить основную литературу, ознакомиться с дополнительной литературой, новыми публикациями в периодических изданиях: журналах, газетах и т.д. При этом учесть рекомендации преподавателя и требования учебной программы. Дорабатывать свой конспект лекции, делая в нем соответствующие записи из литературы, рекомендованной преподавателем и предусмотренной учебной программой. Подготовить тезисы для выступлений по всем учебным вопросам, выносимым на семинар. Готовясь к докладу или реферативному сообщению, обращаться за методической помощью к преподавателю. Составить план-конспект своего выступления. Продумать примеры с целью обеспечения тесной связи изучаемой теории с реальной жизнью. Своевременное и качественное выполнение самостоятельной работы базируется на соблюдении настоящих рекомендаций и изучении рекомендованной литературы.

Самостоятельная работа. Основными способами самостоятельной работы по изучению дисциплины являются: чтение учебников, учебно-методических пособий и другой учебной литературы; регулярное чтение журналов, газет, просмотр и прослушивание теле- и радиопередач; подготовка докладов, научных сообщений и выступление с ними на практических занятиях, научных (научно-практических) конференциях; подготовка и написание рефератов по темам изучаемой дисциплины; выполнение заданий, рекомендованных (заданных) преподавателем; подготовка к зачету

Лабораторная работа. В ходе выполнения лабораторных работ следует изучить основную литературу, ознакомиться с дополнительной литературой, новыми публикациями в периодических изданиях: журналах, газетах и т.д. При этом учесть рекомендации преподавателя и требования учебной программы. Дорабатывать свой конспект лекции, делая в нем соответствующие записи из литературы, рекомендованной преподавателем и предусмотренной учебной

программой. Четко выполнять последовательность операций, оформлять в письменном виде полученные в ходе анализа результаты.

Зачет. Готовиться к зачету необходимо последовательно, с учетом контрольных вопросов, разработанных преподавателем. Сначала следует определить место каждого контрольного вопроса в соответствующем разделе темы учебной программы, а затем внимательно прочитать и осмыслить рекомендованные научные работы, соответствующие разделы рекомендованных учебников. При этом полезно делать хотя бы самые краткие выписки и заметки. Работу над темой можно считать завершенной, если вы сможете ответить на все контрольные вопросы и дать определение понятий по изучаемой теме.

2. РАСПРЕДЕЛЕНИЕ КУРСА ПО ТЕМАМ И ВИДАМ РАБОТ

N	Разделы дисциплины / модуля	С е м е с т р	Виды и часы контактной работы, их трудоемкость (в часах)			Самостоятельная работа
			Лекции	Практические занятия	Лабораторные работы	
1	Тема 1. Основы прикладной социологии, понятие измерения в социологии	2	2	4	0	5
2	Тема 2. Этапы анализа данных. Программное обеспечение в структуре социологического исследования	2	2	4	0	5
3	Тема 3. Обработка социологической информации с помощью Excel	2	4	6	0	10
4	Тема 4. Обработка социологической информации с помощью SPSS. Знакомство с пользовательским интерфейсом программы SPSS	2	4	8	0	10
5	Тема 5. Ввод и редактирование данных в SPSS. Одномерные частотные распределения в SPSS. Двумерные распределения в SPSS и анализ корреляций	2	4	8	0	20
6	Тема 6. Работа с социологической информацией в сети Интернет	2	2	6	0	4
	Итого		18	36	0	54

Форма итогового контроля: зачет во 2 семестре

3. СОДЕРЖАНИЕ КУРСА

Тема 1. Основы прикладной социологии, понятие измерения в социологии

Измерение в социологии. Виды переменных. Организация разных видов переменных. Номинальная, порядковая, количественная шкалы. Операции, допустимые в различных шкалах. Процедуры операциональной интерпретации основных понятий исследования данных. Способы обеспечения требований релевантности, надежности, точности, устойчивости и валидности данных.

Социальная статистика как наука, ее объект и предмет. Социальная статистика как инструмент анализа массовых социальных явлений. Специфика социальной статистики, ее связь с другими науками. Социальная статистика и математика. Социальная статистика и экономика. Социальная статистика и демография. Социальная статистика и социология.

Классификация методов статистического исследования. Статистическое наблюдение. Структура социальной статистики. Социальная статистика и отраслевые статистики. Основные функции социальной статистики: познавательная, прогностическая, дескриптивная, информационная, просветительская. Актуальные проблемы и основные задачи социальной статистики в современных условиях. Организация статистики в Российской Федерации. Структура органов государственной статистики, их основные функции. Государственная и ведомственная статистика. Взаимодействие с международными статистическими службами. Статистические публикации и статистическая периодика. Понятие статистического показателя, его значение с статистическом исследовании. Функции статистических показателей: познавательная, управленческая (контрольно-организационная), стимулирующая. Основные группы статистических показателей: 1) количественные, качественные; 2) индивидуальные, характеризующие единичные процессы, обобщающие; 3) интервальные, моментные. Основные

виды статистических показателей: абсолютные, относительные, средние величины.

Система статистических показателей: показатели населения и трудовых ресурсов; показатели образа жизни населения: показатели уровня жизни (материального, культурного, образования, социального обеспечения, торгового, бытового, жилищно-коммунального обслуживания, здоровья, заболеваемости, обеспеченности учреждениями здравоохранения, спортивными сооружениями, охвата населения физкультурой и спортом и т.д.); показатели национального богатства; показатели производства и обращения общественного продукта; показатели национального дохода и финансов; показатели по труду; баланс народного хозяйства.

Понятие измерения. Аддиктивность и изоморфизм измерения. Соотношение эмпирических и символических систем. Измерение и процедура операционализации основных понятий исследования.

Тема 2. Этапы анализа данных. Программное обеспечение в структуре социологического исследования

Этапы анализа данных. Основные виды анализа данных в социологии: группировка, типологизация, поиск взаимосвязей между переменными, построение индексов, кластерный и факторный анализ, корреляционный анализ, регрессионный анализ. Программное обеспечение в структуре социологического исследования. Применение различных видов программного обеспечения на разных этапах исследования: Excel, SPSS, STATA. Критерии выбора программного обеспечения: специфика данных, наличие необходимых процедур анализа, финансовые возможности или возможности доступа, совместимость данных.

Статистические методы и возможности научного исследования для социолога. Современные методы статистической обработки и анализа данных, полезные для социологических исследований, специализированные цифровые

продукты позволяющие оптимизировать процесс статистического анализа, связь статистического анализа, больших данных и социологических исследований.

Многообразие методов, применяемых социологами для анализа данных. Принципы отбора методов – характер исследовательских гипотез, вопросы исследования, тип данных. Если целью является описание одной характеристики выборки в определенный момент времени, целесообразно ограничиться одномерным анализом. Разнообразные техники многомерного анализа позволяют одновременно исследовать взаимоотношения двух и более переменных и в той или иной форме проверять гипотезы о причинных связях между ними.

Для того, чтобы данные конкретного социологического исследования могли быть статистически обработаны тип данных и особенности их измерения должны быть определены уже на этапе эмпирической и операциональной интерпретации понятий исследования, до начала сбора данных (в первичном исследовании). Это же правило лежит в основе грамотного и успешного статистического анализа. В случае использования вторичных данных набор доступных статистических операций выбирается исходя из типа данных и технических возможностей реализации анализа.

Обработка больших объёмов информации невозможна без специализированных информационных продуктов, так как построение таблиц с тысячами и десятками тысяч наблюдений не целесообразны и физически не возможны. Для решения этой задачи используются статистические пакеты программ (далее СПП) и (или) языки программирования.

Это позволяет справляться основными классами задач описательной и аналитической статистики. Таким образом, использование статистических методов в процессе проведения социологического исследования позволяет достичь многих практических целей, а специализированные программные

продукты позволяют: работать с большим объёмом данных, ускорять процесс описания и анализа, автоматизировать статистические операции, визуализировать результаты

Специализированные цифровые продукты позволяющие оптимизировать процесс статистического анализа. Потребность в средствах статистического анализа данных в различных областях деятельности, особенно в науке, очень велика, что и послужило причиной развития рынка компьютерных программ для статистической обработки данных. Активное развитие получили компьютерные программы, позволяющие проводить статистический анализ больших объемов данных с целью выявления закономерностей, сравнения вероятных альтернатив выбора, построения прогнозов развития событий, обнаружения связей между явлениями и процессами и пр.

Перед пользователями различных категорий встает вопрос выбора оптимального статистического пакета для решения собственных исследовательских задач.

При выборе пакета учитываются следующие параметры:

- соответствие характеру решаемых задач;
- объем обрабатываемых данных;
- требования, предъявляемые к квалификации пользователя (уровень знаний в области статистики);
- имеющееся в наличии компьютерное оборудование;
- стоимость продукта

Тема 3. Обработка социологической информации с помощью Excel

Обработка социологической информации с помощью Excel. Ввод данных. Виды и способы построения переменных. Описательная статистика в Excel. Анализ статистических взаимосвязей: двумерное распределение, описание

группы. Графическое отображение информации. Построение графиков, диаграмм различных видов и инфографики.

Самой доступной, но не специализированной программой для статистической обработки данных является приложение MS Excel. Продукт входит в базовый пакет Microsoft Office. Причинами популярности продукта являются дружелюбный и достаточно интуитивный интерфейс, переведенный на русский язык, тесная интеграция с другими приложениями семейства MS Office. Программа проста в освоении по сравнению с другими цифровыми, виртуальными средствами статистической обработки данных.

Приложение специализировано для работы с электронными таблицами, и, по сути, MS Excel это программа, где сложные математические расчеты являются дополнительными функциями и реализованы в виде встроенных формул. Отрицательной чертой является тот факт, что вычисления, сделанные при помощи базовой версии MS Excel, не признаются научными кругами и печатными изданиями. Без подключения дополнительных модулей в программе невозможно сразу построить качественные наглядные диаграммы и графики. Однако, для полноценного решения сложных статистических и инженерных задач, в Microsoft Excel возможно установить дополнительный набор средств анализа данных (так называемый пакет анализа). С помощью этих инструментов следует задать исходные данные и выбрать необходимые параметры, после чего анализ будет выполнен с помощью подходящей статистической или инженерной макрофункции, а результат будет помещен в назначенный выходной диапазон. Также в форме дополнительных модулей доступны специальные средства, которые позволяют интерпретировать результаты анализа в графическом виде. Для MS Excel существует также макрос-дополнение XLSTAT-Pro, включающее конструктор из более чем 50 статистических функций, которых в большинстве случаев достаточно для рядового применения в анализе массива данных. Данная надстройка является платным дополнением к MS Excel.

Существует и возможность автоматизации расчетов и визуализации данных с помощью языка программирования VBA. Также на рынке присутствует коммерческий пакет XLStat, позволяющий реализовать основные процедуры статистического анализа, доступные в универсальных программах, прямо в Excel.

Тема 4. Обработка социологической информации с помощью SPSS. Знакомство с пользовательским интерфейсом программы SPSS

Обработка социологической информации с помощью SPSS. Знакомство с пользовательским интерфейсом программы SPSS. Главное меню и диалоговые окна. Панели инструментов. Окно редактора данных, окно синтаксиса и окно вывода. Функциональная и пользовательская настройка (шрифты, стили вывода результатов и т.д.) Окно описания переменных. Открытие уже существующего файла с данными. Возможность конвертации (импорта) файлов с данными, созданных в других программах. Сохранение файла данных в различных форматах (экспорт данных). Использование результатов в приложениях WORD, EXCEL, вставка и копирование объектов.

Пакет SPSS (Statistical Package for the Social Sciences) – универсальный статистический пакет компании SPSS Inc. SPSS является модульной программой. Ее основу составляет базовый модуль (SPSS Base), позволяющий осуществлять управление данными и содержащий наиболее распространенные методы статистического анализа данных: проведение описательной статистики; построение линейных и нелинейных моделей; осуществление преобразования данных; проведение факторного, кластерного, дисперсионного анализов; вычисление корреляций; построение графиков; подготовка отчетов и пр.

Для проведения расширенного и углубленного анализа данных могут быть установлены дополнительные модули пакета. Например, модуль IBM SPSS Advanced Statistics предназначен для проведения анализа сложных взаимосвязей при помощи процедур, учитывающих свойства исследуемых

данных, что позволяет продвинуться за рамки базового анализа данных. В модуль встроены мощные инструменты построения моделей. Модуль IBM SPSS Bootstrapping ("Самогенерация") позволяет аналитикам проверять устойчивость построенных моделей, а модуль IBM SPSS Direct Marketing ("Прямой маркетинг") предоставляет возможность маркетологам самостоятельно выполнять основные виды анализа. Модуль IBM SPSS Data Entry автоматизирует процесс разработки анкеты и ввода результатов опросов.

Новые модули пакета также позволяют собирать, структурировать и анализировать данные веб-пространства Программные продукты SPSS, которые интегрируются с решениями для работы с большими данными: SPSS Modeler, SPSS Analytic Server, SPSS Collaboration and Deployment Services, SPSS Analytic Catalyst. Есть интеграция с R и Python

Достоинства SPSS:

- развитый аппарат статистического анализа;
- возможность использовать командный синтаксис (syntax) spss (является своего рода языком программирования). Имеет ряд преимуществ: работа с повторяющимися операциями на обновлённом либо новом массиве данных, работа с учётом операций, требующих изменений, создание и применение пользовательских функций, возможность использования циклов для объёмных массивов данных, более широкие возможности для работы со сложными операциями, касающимися трансформации данных, наличие легко читаемого журнала с комментариями
- универсальность (может быть использован для решения широкого круга вопросов из различных предметных областей, требующих проведения статистического анализа данных);
- широкий набор статистических и графических процедур (более 50 типов диаграмм) анализа данных, а также процедур создания отчетов;

- высокая скорость вычислений, простой и удобный интерфейс;
- детальная контекстно-ориентированная справочная система, позволяющая неопытному пользователю с большей легкостью ориентироваться в программе;
- возможность свободного скачивания демонстрационной версии продукта на официальном сайте компании, наличие версий продукта на различных языках, в том числе русском;
- совместимость с операционными системами windows, mac, linux;
- наличие значительного количества литературы по работе с пакетом.

Недостатки SPSS:

- высокие требования к системе компьютера (требуется минимум 1GB оперативной памяти, 800MB памяти на жестком диске и процессор с частотой 1GHz и выше);
- высокая цена по сравнению со статистическими пакетами аналогичного уровня.

Последняя версия SPSS Statistics V26. В новый выпуск добавлены такие возможности, как квантильная регрессия, ROC-анализ и многие другие функции, которые помогут улучшить результаты анализа.

Тема 5. Ввод и редактирование данных в SPSS. Одномерные частотные распределения в SPSS. Двумерные распределения в SPSS и анализ корреляций

Ввод и редактирование данных в SPSS. Создание новой базы данных. Наблюдение и переменная. Создание имен переменных, создание метки переменных и метки значений, использование специальных кодов для пропущенных значений. Ввод данных (числовых и нечисловых), ввод данных в выбранную область. Изменение типа данных, возможность вырезать,

копировать и вставлять значение данных, добавление и удаление наблюдений, добавление и удаление переменных, изменение порядка переменных, изменение описания переменных. Поиск переменных, поиск наблюдений, поиск значений данных. Одномерные частотные распределения в SPSS. Методы построения одномерных распределений. Получение значений частот и статистик, частотные статистики, диаграммы частот, работа с таблицами. Двумерные частотные распределения в SPSS. Построение таблиц сопряженности, слои таблиц сопряженности, пучковые столбиковые диаграммы для таблиц сопряженности, содержание клеток таблиц, статистики, вычисляемые для таблиц сопряженности, формат таблиц.

Набор статистических методов анализа, который включен в пакет: описательная статистика (базовые статистические методы, проверка нормальности распределения данных); дисперсионный анализ; непараметрическая статистика (анализ таблиц сопряженности, непараметрические сравнения, дисперсионный анализ); контроль качества; анализ выживаемости; кластерный анализ; факторный анализ; дискриминантный анализ; регрессионный анализ; обработка данных (сортировка, отбор, трансформация данных).

Тема 6. Работа с социологической информацией в сети Интернет

Социологическая информация в Интернет: электронные библиотеки, электронные базы данных. Единый архив социологических данных: история, организация. Вторичный анализ данных в режиме on-line, анализ контента социальных сетей, работа с данными органов государственной статистики, работа с готовыми массивами данных социологических и маркетинговых опросов.

Сегодня социологи все чаще используют большие данные для анализа социальных институтов, практик, норм и т.д. Мы будем использовать в связке с терминами «большие» и «открытые» данные, под ними понимаются

машиночитаемые наборы данных, которые выкладываются в сеть государственными, муниципальными, общественными структурами, сообществами гражданских интернет-активистов. Проблемность ситуации заключается не только в возрастающем объеме новых данных, но и в ограниченности доступа к ним. Исследователь ежедневно может видеть в Интернете тысячи релевантных ссылок, открытые машиночитаемые данные, данные, интегрированные в код страниц сайтов. Все они доступны, но требуют такой кропотливой ручной работы по извлечению и анализу. На этом этапе и появляется потребность в развитии новых компетенций социолога таких как умение добывать большие объемы данных, структурировать их в форму поддающуюся анализу обрабатывать их и предоставлять аналитические результаты.

Для этого исследователю необходимо овладеть навыками программирования и работы в статистических средах, с интерфейсом командной строки. Несколько строчек программного кода могут решать сложные, с точки зрения обработки больших массивов данных, задачи

Работа с большими и открытыми данными ведет к необходимости технического перевооружения исследователя. Одно из его направлений — переход со статистических пакетов с «кнопочным» интерфейсом на аналитические инструменты, основанные на интерфейсе командной строки, предполагающей наличие у ученого знаний кодинга и программирования. В качестве привлекательного и популярного аналитического инструмента можно назвать статистическую среду и язык программирования R.

Эта аналитическая платформа распространяется бесплатно и имеет много приверженцев, постоянно совершенствующих как ее программную оболочку, так и количество и объем доступных методов анализа данных. Основные преимущества R:

- большое количество инструментов по адаптации и преобразованию данных под запрос исследователя;
- возможность видеть результаты манипуляций с данными, которые осуществляют те или иные статистические алгоритмы;
- возможность настраивать среду для решения конкретных специфических задач в той или иной области исследований;
- большое число модулей и пакетов, позволяющих использовать самые современные методы анализа данных;
- возможность самостоятельно создавать скрипты, модули, пакеты, позволяющие решать ту или иную задачу в сфере обработки и анализа данных;
- возможность напрямую считывать машиночитаемые данные с Интернет-ресурсов и работать с ними, как с обычными данными;
- большая библиотека графических методов, позволяющих упаковывать данные в привлекательные и легко интерпретируемые рисунки и диаграммы;
- интеграция с текстовыми и издательскими системами, такими, как LaTeX, Microsoft Word, позволяющая создавать аналитические отчеты по итогам исследования непосредственно в среде R, используя ее язык программирования.

- интеграция с Google Analytics, Яндекс.Метрикой

Недостатки R

- R — узкоспециализированный язык
- для эффективного решения задач требуется знание расширений и дополнений

Как показывает практика работы с открытыми данными (и в машиночитаемых форматах, и в просто встроенных в веб-страницы),

исследователь может столкнуться со спектром проблем по управлению данными, которые невозможно решить средствами R. Возникает потребность не просто в освоении статистических платформ, имеющих встроенный язык программирования, а в использовании специализированного языка, благодаря которому можно извлекать и трансформировать данные.

Одним из наиболее приемлемым для решения социологических задач является язык Python.

Основные преимущества Python

- язык высокого уровня, позволяющий писать простые, но функциональные и эффективные скрипты по адаптации и преобразованию данных.
- С помощью Python легко работать с онлайн-ресурсами.
- Ключевое преимущество данного языка программирования — понятный и легко воспринимаемый синтаксис, простота освоения.
- у Facebook есть открытая библиотека для Python.(<https://facebook-sdk.readthedocs.io/en/latest/>)
- программная библиотека на языке Python - Pandas. Он включает в себя огромный функционал для различных манипуляций с данными: статистическая обработка, очистка, трансформация, агрегация. Pandas может стать достойной заменой Excel, если вы много работаете с обработкой массивов данных.
- Python обладает большими возможностями и для визуализации данных. Можно начать с освоения библиотеки Vokeh или Chartify.

Недостатки Python

- - сложнее для изучения

- - для эффективного решения задач требуется знание расширений и дополнений

Стоит отметить, что при помощи R и Python возможно решение и статистических задач «классических исследований», например данных опросов и т.д.

4. ПЛАНЫ СЕМИНАРСКИХ ЗАНЯТИЙ

Тема 1. Основы прикладной социологии, понятие измерения в социологии

1. Понятие «изменение», «данные», «наблюдение», «шкала» в социологии.
2. Виды переменных. Организация разных видов переменных.
3. Номинальная, порядковая, метрические шкалы.
4. Операции, допустимые в различных шкалах.
5. Дискретные не прерывные данные

Литература:

Основная

1. Кулаков А. П. Измерение в социологии : учеб. пособие / А. П. Кулаков ; Новосиб. гос. архитектур.-строит. ун-т. – Новосибирск : НГАСУ (Сибстрин), 2005. – 124 с.
2. Миркин, Б.Г. Введение в анализ данных. Учебник и практикум / Б.Г. Миркин. — Люберцы: Юрайт, 2016. – 174 с.
3. Горяинова, Е.Р. Прикладные методы анализа статистических данных: Учебное пособие / Е.Р. Горяинова, А.Р. Панков, Е.Н. Платонов. — М.: ИД ГУ ВШЭ, 2018. – 310 с.(с.59-113)
4. Толстова Ю.Н. Измерение в социологии: учебное пособие./Ю.Н. Толстова –М.: КДУ, 2007. – 288 с.
5. Кулаичев, А. П. Методы и средства комплексного статистического анализа данных: учебное пособие / А.П. Кулаичев. — 5-е изд., перераб. и доп. — Москва : ИНФРА-М, 2018. — 484 с. — (Высшее образование: Бакалавриат). — www.dx.doi.org/10.12737/25093. – ISBN 978-5-16-012834-4. - Текст: электронный. - URL: <https://znanium.com/catalog/product/975598> (дата обращения: 03.03.2020). – Режим доступа: по подписке.
6. Хорольский, В. В. Медийность текстов в средствах массовой коммуникации : монография / В. В. Хорольский. - 3-е изд., стер. - Москва:

ФЛИНТА, 2020. - 243 с. - ISBN 978-5-9765-2468-2. - Текст: электронный. - URL: <https://znanium.com/catalog/product/1150925> (дата обращения: 03.03.2020). – Режим доступа: по подписке.

7. Ковалев, А.И. Диагностика качества функционирования предприятия : монография / А.И. Ковалев. - Москва ; Вологда : Инфра-Инженерия, 2019. - 400 с. - ISBN 978-5-9729-0328-3. - Текст: электронный. - URL: <https://znanium.com/catalog/product/1053383> (дата обращения: 03.03.2020). – Режим доступа: по подписке.

Дополнительная

1. Волкова, П. А. Статистическая обработка данных в учебно-исследовательских работах : учебное пособие / П.А. Волкова, А.Б. Шипунов. — Москва : ФОРУМ : ИНФРА-М, 2020. — 96 с. — (Высшее образование: Бакалавриат). - ISBN 978-5-00091-710-7. - Текст: электронный. - URL: <https://znanium.com/catalog/product/1091712> (дата обращения: 03.03.2020). – Режим доступа: по подписке.

2. Новиков, А. И. Математические методы в психологии: учебное пособие / А.И. Новиков, Н.В. Новикова, - 2-е изд. - Москва: НИЦ ИНФРА-М, 2018. - 256 с.: - (Высшее образование: Бакалавриат). - ISBN 978-5-16-011253-4. - Текст : электронный. - URL: <https://znanium.com/catalog/product/970004> (дата обращения: 03.03.2020). – Режим доступа: по подписке.

3. Григорьев, А. А. Методы и алгоритмы обработки данных : учебное пособие / А.А. Григорьев. — Москва : ИНФРА-М, 2018. — 256 с. — (Высшее образование: Бакалавриат). — www.dx.doi.org/10.12737/22119. - ISBN 978-5-16-011916-8. - Текст: электронный. - URL: <https://znanium.com/catalog/product/922736> (дата обращения: 03.03.2020). – Режим доступа: по подписке.

Глоссарий:

Измерение в социологическом исследовании – это процедура, с помощью которой измеряемый объект сравнивается с некоторым эталоном и получает числовое выражение в определенном масштабе и шкале.

Переменные – некоторое общее для всех изучаемых объектов, например людей, свойство, конкретные проявления которого могут меняться от объекта к объекту.

Данные – это результаты наблюдений, испытаний, накапливаемые с целью последующего изучения и анализа.

Шкалы – правило, определяющее, каким образом в процессе измерения каждому изучаемому объекту ставится в соответствие некоторое число или другой математический конструкт.

Тема 2. Этапы анализа данных. Программное обеспечение в структуре социологического исследования

1. Группировка, типологизация данных
2. Поиск взаимосвязей между переменными
3. Построение индексов
4. Кластерный, факторный анализ и корреляционный анализ
5. Программное обеспечение в структуре социологического исследования.
6. Критерии выбора программного обеспечения.

Литература:

Основная

1. Кулаичев, А. П. Методы и средства комплексного статистического анализа данных: учебное пособие / А.П. Кулаичев. — 5-е изд., перераб. и доп. — Москва : ИНФРА-М, 2018. — 484 с. — (Высшее образование: Бакалавриат). — www.dx.doi.org/10.12737/25093. - ISBN 978-5-16-012834-4. – Текст: электронный. - URL: <https://znanium.com/catalog/product/975598> (дата обращения: 03.03.2020). – Режим доступа: по подписке.
2. Кацко, И.А. Практикум по анализу данных на компьютере / И.А. Кацко, Н.Б. Паклин. — М.: КолосС, 2017. — 278 с. (с.20-78)
3. Боровиков, В.П. Популярное введение в современный анализ данных в системе STATISTICA: Учебное пособие для вузов / В.П. Боровиков. — М.: Гор. линия-Телеком, 2018. — 288 с. (Главы 4 и 5)
4. Воскобойников, Ю.Е. Регрессионный анализ данных в пакете MATHCAD + CD / Ю.Е. Воскобойников. — СПб.: Лань, 2015. — 224 с. (с. 6-50)
5. Горяинова, Е.Р. Прикладные методы анализа статистических данных: Учебное пособие / Е.Р. Горяинова, А.Р. Панков, Е.Н. Платонов. — М.: ИД ГУ ВШЭ, 2018. — 310 с.(с.59-113)
6. Статистический анализ данных, моделирование и исследование вероятностных закономерностей. Компьютерный подход / Б.Ю.Лемешко, С.Б.Лемешко, С.Н.Постовалов, Е.В.Чимитова; ЭБС Znanium. - М.: НИЦ ИНФРА-М, 2015. - 890 с. (С. 400-462)

Дополнительная

1. Цыпин Александр Павлович, Сорокин Александр Сергеевич Статистические пакеты программ в социально-экономических исследованиях // АНИ: экономика и управление. 2016. №4 (17). С. 379-384.
2. Гржибовский А.М., Иванов С.В., Горбатова М.А. Корреляционный анализ данных с использованием программного обеспечения Statistica и SPSS // Наука и здравоохранение. 2017. №1. С. 7-36

3. Гржибовский А.М., Иванов С.В., Горбатова М.А. Однофакторный линейный регрессионный анализ с использованием программного обеспечения Statistica и SPSS // Наука и здравоохранение. 2017. №2. С. 5-33

4. IBM SPSS Statistics 20 и Amos: Профессиональный статистический анализ данных. Практическое руководство. А. Наследов. С-Пб.: Питер, 2013, - 416 с. (С. 113-149)

Глоссарий:

Программное обеспечение – компьютерное программное обеспечение для обработки эмпирических данных. В процессе анализа данных социолог использует различные программы: текстовый редактор (для подготовки отчета), пакет иллюстративной компьютерной графики (для подготовки графиков в отчет или слайдов для публичного выступления результатам) и, конечно, пакет программ для анализа данных.

Корреляция (корреляционная зависимость) – это некая зависимость двух или более случайных величин

Индекс – агрегированный количественный показатель, обобщающий первичную социологическую информацию, полученную в ходе измерения с помощью одной или нескольких шкал

Тема 3. Обработка социологической информации с помощью Excel

1. Ввод данных
2. Виды и способы построения переменных
3. Описательная статистика в Excel
4. Анализ статистических взаимосвязей: двумерное распределение, описание группы.
5. Графическое отображение информации – построение графиков и диаграмм различных видов.

Литература:

Основная

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> (дата обращения: 03.03.2020). – Режим доступа: по подписке.

2. Форман, Д. Много цифр. Анализ больших данных при помощи Excel / Форман Д.; Пер. с англ. Соколовой А. - Москва : Альпина Паблишер, 2016. - 461 с. ISBN 978-5-9614-5032-3. - Текст: электронный. - URL: <https://znanium.com/catalog/product/551044> (дата обращения: 03.03.2020). – Режим доступа: по подписке.

3. Горяинова, Е.Р. Прикладные методы анализа статистических данных: Учебное пособие / Е.Р. Горяинова, А.Р. Панков, Е.Н. Платонов. — М.: ИД ГУ ВШЭ, 2018. — 310 с.

Дополнительная

1. Волкова, П. А. Статистическая обработка данных в учебно-исследовательских работах : учебное пособие / П.А. Волкова, А.Б. Шипунов. — Москва : ФОРУМ : ИНФРА-М, 2020. — 96 с. — (Высшее образование: Бакалавриат). - ISBN 978-5-00091-710-7. - Текст: электронный. - URL: <https://znanium.com/catalog/product/1091712> (дата обращения: 03.03.2020). – Режим доступа: по подписке.

2. Новиков, А. И. Математические методы в психологии: учебное пособие / А.И. Новиков, Н.В. Новикова, - 2-е изд. - Москва: НИЦ ИНФРА-М, 2018. - 256 с.: - (Высшее образование: Бакалавриат). - ISBN 978-5-16-011253-4. - Текст : электронный. - URL: <https://znanium.com/catalog/product/970004> (дата обращения: 03.03.2020). – Режим доступа: по подписке.

3. Григорьев, А. А. Методы и алгоритмы обработки данных : учебное пособие / А.А. Григорьев. — Москва : ИНФРА-М, 2018. — 256 с. — (Высшее образование: Бакалавриат). — www.dx.doi.org/10.12737/22119. - ISBN 978-5-16-011916-8. - Текст: электронный. - URL:

<https://znanium.com/catalog/product/922736> (дата обращения: 03.03.2020). –
Режим доступа: по подписке.

Глоссарий:

Описательная статистика – занимается обработкой эмпирических данных, их систематизацией, наглядным представлением в форме графиков и таблиц, а также их количественным описанием посредством основных статистических показателей.

Распределение – метод описания данных, позволяющий определить характерные свойства и закономерности изменения признаков изучаемой совокупности явлений.

Тема 4. Обработка социологической информации с помощью SPSS.

Знакомство с пользовательским интерфейсом программы SPSS

1. Интерфейс программы SPSS: главное меню, диалоговые окна, панели инструментов.
2. Окно редактора данных, окно синтаксиса и окно вывода.
3. Открытие готовых наборов данных и создание новых
4. . Возможности импорта и экспорта данных.

Литература:

Основная

1. Крыштановский, А.О. Анализ социологических данных с помощью пакета SPSS: учебное пособие / А.О.Крыштановский. - М.: Изд. дом ГУ ВШЭ, 2006. - 281 с.
2. Наследов, А.Д. IBM SPSS Statistics 20 и AMOS: профессиональный статистический анализ данных / А.Д. Наследов. — СПб.: Питер, 2018. — 416 с.

3. Кулаичев, А. П. Методы и средства комплексного статистического анализа данных: учебное пособие / А.П. Кулаичев. — 5-е изд., перераб. и доп. — Москва : ИНФРА-М, 2018. — 484 с. — (Высшее образование: Бакалавриат). — www.dx.doi.org/10.12737/25093. - ISBN 978-5-16-012834-4. - Текст: электронный. - URL: <https://znanium.com/catalog/product/975598> (дата обращения: 03.03.2020). – Режим доступа: по подписке.

Дополнительная

1. Наследов, А.Д. IBM SPSS Statistics 20 и AMOS: профессиональный статистический анализ данных / А.Д. Наследов. — СПб.: Питер, 2018. — 416 с. Кацко, И.А. Практикум по анализу данных на компьютере: учебно-практическое пособие / И.А.Кацко, Н.Б.Паклин; под ред. проф. Г.В.Гореловой. - М.: КолосС, 2009. - 278 с. Гриф МО РФ б. Malhotra, N.K. Basic Marketing Research / N.K.Malhotra. - 4th ed. - Edinbughr: Pearson Education Limited, 2014. - 670 p.
2. Лагутин, М.Б. Наглядная математическая статистика: учебное пособие / М.Б.Лагутин. - 3-е изд; испр. - М.: БИНОМ, Лаборатория знаний, 2013. - 472 с.
3. Статистические методы анализа данных [Электронный ресурс]: учебник / Л.И. Ниворожкина, С.В. Арженовский, А.А. Рудяга [и др.]; под общ. ред. д-ра экон. наук, проф. Л.И. Ниворожкиной; ЭБС Znanium. — М.: РИОР: ИНФРА-М, 2016. — 333 с.

Глоссарий:

Интерфейс — это «проводник» между человеком и программой, операционной системой, техническим устройством или способ взаимодействия приложений между собой. Человек дает команды с помощью интерфейса, устройство их анализирует и отвечает.

Экспорт и импорт данных – это автоматический или полуавтоматический ввод и вывод наборов данных между различными программными приложениями.

Тема 5. Ввод и редактирование данных в SPSS. Одномерные частотные распределения в SPSS. Двумерные распределения в SPSS и анализ корреляций

1. Создание базы социологических данных в SPSS. Создание новой базы данных.
2. Кодирование переменных, принцип категорий и принцип дихотомий.
3. Ввод информации о переменных – имен переменных, создание метки переменных и метки значений, использование специальных кодов для пропущенных значений.
4. Одномерные частотные распределения в SPSS. Методы построения одномерных распределений.
5. Двумерные частотные распределения в SPSS. Построение таблиц сопряженности, строки таблиц сопряженности, статистики, вычисляемые для таблиц сопряженности.

Литература:

Основная

1. Крыштановский, А.О. Анализ социологических данных с помощью пакета SPSS: учебное пособие / А.О.Крыштановский. - М.: Изд. дом ГУ ВШЭ, 2006. - 281 с.
2. Наследов, А.Д. IBM SPSS Statistics 20 и AMOS: профессиональный статистический анализ данных / А.Д. Наследов. — СПб.: Питер, 2018. — 416 с.
3. Кулаичев, А. П. Методы и средства комплексного статистического анализа данных: учебное пособие / А.П. Кулаичев. — 5-е изд., перераб. и доп. — Москва : ИНФРА-М, 2018. — 484 с. — (Высшее образование: Бакалавриат).

— www.dx.doi.org/10.12737/25093. - ISBN 978-5-16-012834-4. - Текст: электронный. - URL: <https://znanium.com/catalog/product/975598> (дата обращения: 03.03.2020). – Режим доступа: по подписке.

4. Бююль А., Цефель П. SPSS: искусство обработки информации. Анализ статистических данных и восстановление скрытых закономерностей. М., ДиаСофт, 2005 г. – 608 с.

Дополнительная

1. Горяинова, Е.Р. Прикладные методы анализа статистических данных: Учебное пособие / Е.Р. Горяинова, А.Р. Панков, Е.Н. Платонов. — М.: ИД ГУ ВШЭ, 2018. — 310 с.

2. Гржибовский А.М., Иванов С.В., Горбатова М.А. Корреляционный анализ данных с использованием программного обеспечения Statistica и SPSS // Наука и здравоохранение. 2017. №1. С. 7-36

3. Кацко, И.А. Практикум по анализу данных на компьютере: учебно-практическое пособие / И.А.Кацко, Н.Б.Паклин; под ред. проф. Г.В.Гореловой. - М.: КолосС, 2009. - 278 с. Гриф МО РФ 6. Malhotra, N.K. Basic Marketing Research / N.K.Malhotra. - 4th ed. - Edinbugh: Pearson Education Limited, 2014. - 670 p.

4. Ратникова, Т.А. Анализ панельных данных и данных о длительности состояний: учебное пособие / Т.А.Ратникова, К.К.Фурманов; Нац. исслед. ун-т Высшая школа экономики. - М.: Изд. дом ВШЭ, 2014. - 373 с.

Глоссарий:

Частотное распределение признака – закономерность встречаемости разных значений признака.

Частота – количество наблюдений, в которых признак принимает определенное значение или находится в определенном интервале.

Метка переменной – Используется, когда смысл переменной недостаточно точно отражен в имени переменной. Это поле для ввода полного названия переменной (обычно –номер и формулировка вопроса).

Значения переменной – вводятся коды и их значения обычно –коды ответов на вопрос.

Тема 6. Работа с социологической информацией в сети Интернет

1. Электронные библиотеки, электронные базы данных.
2. Единый архив социологических данных: история, организация.
3. Вторичный анализ данных в режиме on-line, анализ контента социальных сетей, работа с данными органов государственной статистики
4. Большие данные как вид данных и совокупность методов анализа
5. Программные средства с открытым исходным кодом и языки программирования для средство статистического анализа в социологии

Литература:

Основная

1. Кабаков, Р. R в действии. Анализ и визуализация данных в программе R / Р. Кабаков. — М.: ДМК, 2014, 589с. (С. 29-51)
2. Лесковец, Ю. Анализ больших наборов данных / Ю. Лесковец, А. Раджараман. — М.: ДМК, 2016. — 498 с. (С.122-180)
3. Маккинли, У. Python и анализ данных / У. Маккинли. — М.: ДМК, 2015. — 482 с. (С.28-92)
4. О. Б. Божков «Большая социология: расширение пространства данных» // Социологический журнал. 2015. №1.С. 181-184
5. Одинцов Александр Владимирович Открытость баз данных как условие формирования «Больших данных» в социологии // Концепт. 2017. №12. С. 101-106

6. Смирнов В. А. Новые компетенции социолога в эпоху «Больших данных» // Мониторинг. 2015. №2 (125).С. 44-54

7. Катерина Губа Большие данные в социологии: новые данные, новая социология? // Социологическое обозрение. 2018. №1.С.213-236

Дополнительная

1. Люк Д. Анализ сетей (графов) в среде R. Руководство пользователя - М.: ДМК Пресс, 2016. - 248 с.

2. Храмов Д. А. Сбор данных в интернете на языке R. - М.: ДМК Пресс, 2016. - 282 с.

3. Коэльо Л. П., Ричард В. Построение систем машинного обучения на языке Python / пер. с англ. Слинкин А. А. - М.: ДМК Пресс, 2016. - 302 с.

4. Маккинли У. Python и анализ данных / пер. с англ. Слинкин А. А. - М.: ДМК Пресс, 2015. - 482 с.

5. Мастицкий С. Э. Визуализация данных с помощью ggplot2. - М.: ДМК Пресс, 2016. - 222 с.

Глоссарий:

Электронные библиотеки – упорядоченная коллекция разнородных электронных документов, снабжённых средствами навигации и поиска. Может быть веб-сайтом, где постепенно накапливаются различные тексты и медиафайлы, каждый из которых самодостаточен и в любой момент может быть востребован читателем.

Вторичный анализ данных – вид исследования, в котором повторно анализируются данные, собранные и, как правило, опубликованные ранее или/и другими исследователями. Результатом вторичного анализа данных могут быть любые интерпретации, выводы и заключения, отличающиеся от тех, которые были получены и опубликованы авторами.

Большие данные – это структурированные или неструктурированные массивы данных большого объема. Их обрабатывают при помощи специальных

автоматизированных инструментов, чтобы использовать для статистики, анализа, прогнозов и принятия решений.

5. ПРИМЕР ЛАБОРАТОРНОЙ РАБОТЫ

1. Порядок проведения и процедура оценивания

Этапы работы:

1. Уясните цель выполнения работы. Составьте план действий, необходимых для достижения поставленной цели.

2. Проверьте свою подготовленность к выполнению работы, в случае затруднений обращайтесь к теоретическому материалу учебника.

3. Проверьте наличие на вашем лабораторном столе необходимого оборудования.

4. Ознакомьтесь с описанием лабораторной работы. Если возникли сомнения, проконсультируйтесь у преподавателя или лаборанта. Если вопросов нет, приступайте к работе.

5. Используя инструкции произведите необходимые расчёты и операции.

6. Сформулируйте выводы на основании полученных результатов, запишите их в текстовом редакторе.

7. В конце занятия необходимо привести в порядок рабочее место – закрыть все программы, выключить компьютер.

2. Критерии оценивания

Баллы в интервале 86-100% от максимальных ставятся, если обучающийся:

Продемонстрирован высокий уровень знаний и умений, необходимых для выполнения задания. Работа полностью соответствует требованиям профессиональной деятельности. Отличная способность применять имеющиеся знания и умения для решения практических задач. Высокий уровень креативности, самостоятельности. Соответствие выбранных методов поставленным задачам.

Баллы в интервале 71-85% от максимальных ставятся, если обучающийся:

Продемонстрирован средний уровень знаний и умений, необходимых для выполнения задания. Работа в основном соответствует требованиям профессиональной деятельности. Хорошая способность применять имеющиеся знания и умения для решения практических задач. Средний уровень креативности, самостоятельности. Выбранные методы в целом соответствуют поставленным задачам.

Баллы в интервале 56-70% от максимальных ставятся, если обучающийся:

Продемонстрирован низкий уровень знаний и умений, необходимых для выполнения задания. Работа частично соответствует требованиям профессиональной деятельности. Удовлетворительная способность применять имеющиеся знания и умения для решения практических задач. Низкий уровень креативности, самостоятельности. Выбранные методы частично соответствуют поставленным задачам.

Баллы в интервале 0-55% от максимальных ставятся, если обучающийся:

Продемонстрирован неудовлетворительный уровень знаний и умений, необходимых для выполнения задания. Работа не соответствует требованиям профессиональной деятельности. Неудовлетворительная способность применять имеющиеся знания и умения для решения практических задач. Недостаточный уровень креативности, самостоятельности. Выбранные методы не соответствуют поставленным задачам.

3. Содержание оценочного средства

Лабораторная работа по курсу «Анализ социологической информации в статистических программах» включает в себя следующие этапы работ.

- подготовка статистических данных для обработки в SPSS;

- работа с основными блоками и элементами меню SPSS, возможностями построения собственных сценариев на внутреннем языке SPSS, использование графических возможностей пакета.

- применение стандартных статистических функций SPSS к готовому массиву данных и интерпретация полученных результатов;

Для выполнения самостоятельных заданий по теме 2 используется массив данных. Задания

выполняются в программе SPSS.

1. Составить задание на обработку данных (программа, методический инструментарий

предлагаются преподавателем).

2. Составить макеты аналитических таблиц по заданным программным основаниям.

3. Сформулировать задачу для выполнения описательного (дескриптивного) анализа (рассчитать показатели средней тенденции, показатели колебания признака)

4. Рассчитать критерий «хи-квадрат», проверить статистическую гипотезу о независимости признаков.

5. Рассчитать коэффициенты связи в рамках модельной задачи.

6. Рассчитать коэффициенты связи для номинальных данных (коэффициент Крамера V , коэффициент сопряженности C , меру Гудмена-Краскэла).

7. Рассчитать коэффициенты связи для дихотомических данных (четырёхклеточных таблиц сопряженности).

8. Рассчитать коэффициенты ранговой корреляции.

9. Сформулировать задачу для выполнения регрессионного анализа, выбрать переменные, выполнить регрессионный анализ и сделать выводы по полученным результатам.

10. Сформулировать задачу для выполнения факторного анализа, выбрать переменные, выполнить факторный анализ и дать содержательную интерпретацию полученных результатов факторизации.

11. Сформулировать задачу для выполнения кластерного анализа методом к-средних, выбрать переменные, выполнить кластерный анализ и дать содержательную интерпретацию результатов.

6. ВОПРОСЫ К ЗАЧЕТУ

Формулировки зачетных вопросов:

1. Применение различных видов программного обеспечения на разных этапах исследования
2. Критерии выбора программного обеспечения
3. Ввод данных в Excel
4. Описательная статистика в Excel
5. Особенности обработки ограниченных и неограниченных переменных в Excel
6. Анализ статистических взаимосвязей в Excel
7. Создание базы первичных данных в SPSS
8. Виды переменных. Организация разных видов переменных в SPSS
9. Преобразование и отбор информации в SPSS
10. Работа с поливариантными переменными в SPSS
11. Описательная статистика в SPSS
12. Построение одномерных распределений в SPSS
13. Анализ статистических взаимосвязей в SPSS
14. Построение двухмерных распределений в SPSS.
15. Регрессионный анализ в SPSS
16. Многомерный анализ в SPSS
17. Социологическая информация в Интернет
18. Единый архив социологических данных: история, организация работы

- 19.Что такое электронная таблица и каково ее назначение?
- 20.Как обозначаются столбцы и строки в MS EXCEL?
- 21.Как изменить ширину столбца (высоту строки)?
- 22.С какими типами данных работает MS EXCEL?
- 23.Рабочее окно пакета STATISTICA
- 24.Создание файла данных
- 25.Форматирование файла данных
- 26.Сохранение файла данных
- 27.Построение графиков в пакета STATISTICA
- 28.Сопоставление двух параллельных рядов
- 29.Построение аналитической группировки
- 30.Графический метод (построение поля корреляции)

7.ТЕСТЫ ДЛЯ ПРОВЕРКИ ОСТАТОЧНЫХ ЗНАНИЙ

1. Статистика как наука изучает:

- a) единичные явления;
- b) массовые явления;
- c) периодические события.

2. Статистическая совокупность ? это:

- a) множество изучаемых разнородных объектов;
- b) множество единиц изучаемого явления;
- c) группа зафиксированных случайных событий.

3. Основными задачами статистики на современном этапе являются: а) исследование преобразований экономических и социальных процессов в обществе; б) анализ и прогнозирование тенденций развития экономики; в) регламентация и планирование хозяйственных процессов;

- a) а, в
- b) а, б
- c) б, в

4. Статистический показатель дает оценку свойства изучаемого явления:

- a) количественную;
- b) качественную;

с) количественную и качественную.

5. Закон больших чисел утверждает, что:

а) чем больше единиц охвачено статистическим наблюдением, тем лучше проявляется общая закономерность;

б) чем больше единиц охвачено статистическим наблюдением, тем хуже проявляется общая закономерность;

с) чем меньше единиц охвачено статистическим наблюдением, тем лучше проявляется общая закономерность.

6. Статистическое наблюдение ? это:

а) научная организация регистрации информации;

б) оценка и регистрация признаков изучаемой совокупности;

с) работа по сбору массовых первичных данных;

д) обширная программа статистических исследований.

7. Назовите виды статистического наблюдения по степени охвата единиц совокупности:

а) анкета;

б) непосредственное;

с) сплошное;

д) текущее.

8. Является ли статистическим наблюдением наблюдения покупателя за качеством товаров или изменением цен на городских рынках?

- a) да;
- b) нет.

9. Ошибка репрезентативности относится к:

- a) сплошному наблюдению;
- b) не сплошному выборочному наблюдению.

10. Статистическая сводка - это:

- a) систематизация и подсчет итогов зарегистрированных фактов и данных;
- b) форма представления и развития изучаемых явлений;
- c) анализ и прогноз зарегистрированных данных.

11. Статистическая группировка - это:

- a) объединение данных в группы по времени регистрации;
- b) расчленение изучаемой совокупности на группы по существенным признакам;
- c) образование групп зарегистрированной информации по мере ее поступления.

12. Группировочные признаки, которыми одни единицы совокупности обладают, а другие - нет, классифицируются как:

- a) факторные;
- b) атрибутивные;
- c) альтернативные.

13. К каким группировочным признакам относятся: образование сотрудников, профессия бухгалтера, семейное положение:

- a) к атрибутивным
- b) к количественным

14. Ряд распределения - это:

- a) упорядоченное расположение единиц изучаемой совокупности по группам;
- b) ряд значений показателя, расположенных по каким-то правилам.

15. К каким группировочным признакам относятся: сумма издержек обращения, объем продаж, стоимость основных фондов

- a) к дискретным;
- b) к непрерывным.

16. Какие виды статистических таблиц встречаются:

- a) одномерные и перекрестные (простые и комбинационные);
- b) линейные и нелинейные.

16. Статистический показатель - это

- a) размер изучаемого явления в натуральных единицах измерения;
- b) количественная характеристика свойств в единстве с их качественной определенностью;
- c) результат измерения свойств изучаемого объекта.

17. Вычисление средних величин - это

- a) способ изучения структуры однородных элементов совокупности;
- b) прием обобщения индивидуальных значений показателя;
- c) метод анализа факторов.

18. Значения признака, повторяющиеся с наибольшей частотой, называется

- a) модой;
- b) медианой.

19. Медиана в ряду распределения с четным числом членов ряда равна

- a) полусумме двух крайних членов;
- b) полусумме двух срединных членов.

20. Что понимается в статистике под термином ?вариация показателя??

- a) изменение величины показателя;
- b) изменение названия показателя;
- c) изменение размерности показателя.

21. Укажите показатели вариации

- a) мода и медиана;
- b) разброс и дисперсия;
- c) темп роста и прироста.

22. Показатель дисперсии - это:

- a) квадрат среднего отклонения;
- b) средний квадрат отклонений;
- c) отклонение среднего квадрата.

23. Среднеквадратическое отклонение характеризует

- a) взаимосвязь данных;
- b) разброс данных;
- c) динамику данных.

24. Размах вариации исчисляется как

- a) разность между максимальным и минимальным значением показателя;
- b) разность между первым и последним членом ряда распределения.

25. При каком значении линейного коэффициента корреляции связь между Y и X можно признать более существенной:

- a) $r_{yx} = 0,25$;
- b) $r_{yx} = 0,14$;
- c) $r_{yx} = - 0,57$

26. Примером нелинейного мультимедиа являются:

- a) кино
- b) компьютерная игра
- c) слайд-шоу

27. Какое утверждение является неверным?

- a) презентация может представлять собой сочетание текста, гиперссылок, компьютерной анимации, графики, видео, музыки и звукового ряда, которые организованы в единую среду.
- b) презентация имеет сюжет, сценарий и структуру, организованную для удобного восприятия информации.
- c) презентация не предназначена для демонстрации

28. Программой для создания презентаций является

- a) microsoft powerpoint
- b) access
- c) adobe fotoshop

29. Инфографика – это:

- a) графический способ подачи информации, данных и знаний.
- b) информационная подача дидактического материала
- c) информация о графических объектах

30. Анимация используется:

- a) в статичной инфографике
- b) в динамичной инфографике
- c) в печатной инфографике

31. Строка на листе Переменные является в Редакторе данных?

Выберите один ответ:

- a) Объектом, наблюдением
- b) Значения отдельной переменной для объектов
- c) Перечнем свойств отдельной переменной
- d) Свойством переменных

32. Какое окно всегда должно быть открыто при работе с программой SPSS?

Выберите один ответ:

- a) Окно Вывода
- b) Окно Редактора данных
- c) Окно Редактора синтаксиса
- d) Окно Редактора скриптов

33. Каким образом можно корректировать недопустимые значения? (Выберите два правильных ответа)

Выберите несколько ответов:

- a) Удалить наблюдения с недопустимыми значениями из файла данных
- b) Обратиться к исходным источникам информации (анкетам, опросникам)
- c) Перекодировать недопустимые значения в пользовательские пропущенные
- d) Удалить переменные с недопустимыми значениями из файла данных
- e) Выполнить процедуру отбора наблюдений

34. Какими из перечисленных свойств обладают метки значений?

Выберите один ответ:

- a) Первая метка значения в переменной обязательно должна включать в себя формулировку вопроса
- b) Метки значений представляют собой варианты ответа на вопрос, каждому из которых соответствует текстовый или числовой код
- c) Метки значений не могут превышать по длине 56 символов
- d) В тексте меток значений могут быть использованы только числовые символы

35. Чем является Столбец на листе Данные является в Редакторе данных?

Выберите один ответ:

- a) Объектом, наблюдением
- b) Значения отдельной переменной для объектов

- c) Перечень свойств отдельной переменной
- d) Свойство переменных

36. Чем является Строка на листе Данные является в Редакторе данных?

Выберите один ответ:

- a) Объектом, наблюдением
- b) Значением отдельной переменной для объектов
- c) Перечнем свойств отдельной переменной
- d) Свойство переменных

37. Каковы результаты работы процедуры «Перекодировать в другие переменные»? (Выберите два правильных ответа)

- a. Появление в Редакторе данных новой переменной
- b. Внесение изменений в значения исходной переменной
- c. Изменение меток значений исходной переменной
- d. Создание меток значений новой переменной
- e. Создание метки переменной для новой переменной

38. Что не относится к описательным статистикам в SPSS

- a) «среднее»
- b) «сумма»
- c) разброс (меры рассеяния)
- d) все названное относится к описательным статистикам

39. Таблицы сопряженности в SPSS позволяют

- a) Узнать средние значения переменной
- b) 2. Оценить коэффициент сопряженности переменных
- c) Позволяют получить наглядное изображение (в виде частотной таблицы) совместного распределения двух переменных

40. С помощью какой вкладки в SPSS можно вычислить хи-квадрат?

- a) «Таблицы»
- b) «Статистики»
- c) «Данные»

8. РЕКОМЕНДОВАННАЯ ЛИТЕРАТУРА

Основная 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> (дата обращения: 03.03.2020). – Режим доступа: по подписке.
2. Кулаичев, А. П. Методы и средства комплексного статистического анализа данных: учебное пособие / А.П. Кулаичев. — 5-е изд., перераб. и доп. — Москва : ИНФРА-М, 2018. — 484 с. — (Высшее образование: Бакалавриат). — www.dx.doi.org/10.12737/25093. - ISBN 978-5-16-012834-4. - Текст: электронный. - URL: <https://znanium.com/catalog/product/975598> (дата обращения: 03.03.2020). – Режим доступа: по подписке.
3. Форман, Д. Много цифр. Анализ больших данных при помощи Excel / Форман Д.; Пер. с англ. Соколовой А. - Москва : Альпина Пабlishер, 2016. - 461 с. ISBN 978-5-9614-5032-3. - Текст: электронный. - URL: <https://znanium.com/catalog/product/551044> (дата обращения: 03.03.2020). – Режим доступа: по подписке.
4. Хорольский, В. В. Медийность текстов в средствах массовой коммуникации : монография / В. В. Хорольский. - 3-е изд., стер. - Москва: ФЛИНТА, 2020. - 243 с. - ISBN 978-5-9765-2468-2. - Текст: электронный. - URL: <https://znanium.com/catalog/product/1150925> (дата обращения: 03.03.2020). – Режим доступа: по подписке.
5. Ковалев, А.И. Диагностика качества функционирования предприятия : монография / А.И. Ковалев. - Москва ; Вологда : Инфра-Инженерия, 2019. - 400 с. - ISBN 978-5-9729-0328-3. - Текст: электронный. - URL: <https://znanium.com/catalog/product/1053383> (дата обращения: 03.03.2020). – Режим доступа: по подписке.

Дополнительная литература

1. Волкова, П. А. Статистическая обработка данных в учебно-исследовательских работах : учебное пособие / П.А. Волкова, А.Б. Шипунов. — Москва : ФОРУМ : ИНФРА-М, 2020. — 96 с. — (Высшее образование: Бакалавриат). - ISBN 978-5-00091-710-7. - Текст: электронный. - URL: <https://znanium.com/catalog/product/1091712> (дата обращения: 03.03.2020). – Режим доступа: по подписке.

2. Новиков, А. И. Математические методы в психологии: учебное пособие/А.И. Новиков, Н.В. Новикова, - 2-е изд. - Москва: НИЦ ИНФРА-М, 2018. - 256 с.: - (Высшее образование: Бакалавриат). - ISBN 978-5-16-011253-4. - Текст : электронный. - URL: <https://znanium.com/catalog/product/970004> (дата обращения: 03.03.2020). – Режим доступа: по подписке.

3. Григорьев, А. А. Методы и алгоритмы обработки данных : учебное пособие / А.А. Григорьев. — Москва : ИНФРА-М, 2018. — 256 с. — (Высшее образование: Бакалавриат). — www.dx.doi.org/10.12737/22119. - ISBN 978-5-16-011916-8. - Текст: электронный. - URL: <https://znanium.com/catalog/product/922736> (дата обращения: 03.03.2020). – Режим доступа: по подписке.