**Министерство образования и науки Российской Федерации**

**Федеральное государственное автономное образовательное учреждение**

**высшего профессионального образования**

**Казанский (Приволжский) федеральный университет**

Институт социально-философских наук и массовых коммуникаций

Кафедра общей и этнической социологии

Ермолаева П.О.

Компьютерные способы обработки социологической информации

Краткий конспект лекций

Казань, 2014

**Направление:** 031600.62 «Реклама и связи с общественностью» (программа бакалавриата, очное обучение)

**Дисциплина:** «Компьютерные способы обработки социологической информации», 1 курс, очное обучение, форма контроля: экзамен.

**Учебный план:** Профессиональный цикл,очное, 2014г.

**Количество часов:** 72 ч. (в том числе: лекции – 16, практические занятия – 18, самостоятельная работа – 38)

**Аннотация:** курс направлен на освоение основных методов анализа количественной социологической информации. В ходе изучения курса предлагается ознакомить студентов с программой Excel и уделить особое внимание специализированной программе SPSS, являющейся мощной системой статистического анализа и управления данными. В процессе изучения дисциплины студенты познакомятся с характеристиками современных программных средств анализа социологических данных. Студенты познакомятся с базами социологических данных, научатся преобразовывать данные, работать с файлами, изучат простые методики статистического анализа. В ходе практических занятий студенты будут проводить пилотажные социологические исследования в соответствии со сферой своих интересов, формировать базу данных и анализировать массив в программе SPSS на основе полученных теоретических знаний. При этом работа с данными подразумевает не только владение техническими процедурами, но также и возможность грамотно интерпретировать и использовать полученные результаты.

**Темы:**

1. Основы прикладной социологии

2. Измерение в социологии

3. Этапы анализа данных

4. Программное обеспечение в структуре социологического исследования

5. Обработка социологической информации с помощью Еxсel

6. Обработка социологической информации с помощью SPSS. Знакомство с пользовательским интерфейсом программы SPSS

7. Одномерные частотные распределения в SPSS

8. Двумерные частотные распределения в SPSS

9. Работа с социологической информацией в Интернет

**Ключевые слова:** SPSS, Excel, анализ данных, одномерное частотное распределение, двумерное частотно распределение.

**Автор курса:** Ермолаева Полина Олеговна, старший преподаватель кафедры общей и этнической социологии, кандидат социологических наук, доцент, тел.:(843) 2920010, email: polina.ermolaeva@gmail.com

**Дата начала эксплуатации:** 24 ноября 2013 года

**Доступность:** записанные на курс пользователи

**Язык интерфейса:** русский

URL: http://tulpar.kpfu.ru/course/view.php?id=769

### **О Г Л А В Л Е Н И Е**

#### 

Литература 30

Интернет-ресурсы 32

Глоссарий 34

Вопросы для самоконтроля 37

### **Лекционные занятия**

**Тема 1. Основы прикладной социологии**

**Аннотация**. В данной теме раскрываются основы прикладной социологии, техники и методические приемы проведения конкретного социологического исследования

**Ключевые слова***: социологическое исследование, программа социологического исследования, выборочный метод, выборочная совокупность, репрезентативность выборки, методы сбора социологической информации, опрос*

**Методические рекомендации**. Первая тема – основная, от ее успешного освоения будет зависеть понимание студентами роли анализа данных в структуре конкретного социологического исследования.

**Информационные ресурсы.**

Батыгин Г.С. Лекции по методологии социологических исследований: учебник для вузов. – М.: Аспект Пресс, 1995. – 286 с.

Российское общество 2000-х годов. Теоретико-методологические подходы к исследованию. Сборник научных статей. – Казань: Изд-во Казан. техн.ун-та, 2009. – 415 с.

Ядов В.А. Социологическое исследование: методология, программа, методы. – Изд-во «Самарский университет», 1995. – 330 с.

Ядов В.А. Стратегия социологического исследования. Описание, объяснение, понимание социальной реальности. – М.: «Добросвет», 1998. – 395 с.

**Глоссарий**

***социологическое исследование -*** система теоретических и эмпирических процедур, позволяющих получить новые знания о соц. объекте для изучения социальных проблем.

***программа социологического исследования -*** включает весь комплекс методологических и теоретических задач, который решет социолог, а также общий набросок процедур и технических приемов исследования

***выборочный метод***- метод исследования, позволяющий делать заключения о характере распределения изучаемых признаков генеральной совокупности на основании рассмотрения некоторой ее части, а именно выборочной совокупности.

***опрос* -** метод сбора социальной информации об изучаемом объекте в ходе непосредственного (интервью) или опосредованного (анкетирование) общения исследователя и опрашиваемого путем регистрации ответов на вопросы, вытекающие из целей и задач исследования.

***выборочная совокупность*** - часть объектов генеральной совокупности, выступающих в качестве объектов наблюдения.

***репрезентативность выборки*** - свойство выборочной совокупности представлять параметры генеральной совокупности.

**Вопросы по теме:**

1. Социологическое исследование: понятие, виды

2. Программа социологического исследования

3. Формирование выборочной совокупности

4. Общая характеристика методов сбора социологической информации

5. Метод опроса в социологии

6. Методика разработки социологической анкеты

7. Основные принципы построения анкеты

8. Структура опросного листа

Социологическое исследование-система теоретических и эмпирических процедур, позволяющих получить новые знания о соц. объекте для изучения социальных проблем. Типы социологических исследований:

1. По типу логической задачи, которая решается в ходе этого исследования: пилотажные (разведовательное); описательные (дескриптивные); экспериментальные

2.По отношению к объекту исследования: монографические; сравнительное (панельное, трендовое)

3.По срокам проведения: долгосрочные; среднесрочные; краткосрочные; экспресс-исследование

Программа социологического исследования включает весь комплекс методологических и теоретических задач, который решет социолог, а также общий набросок процедур и технических приемов исследования. Состоит из методологического и процедурного разделов.

Опрос - метод сбора социальной информации об изучаемом объекте в ходе непосредственного (интервью) или опосредованного (анкетирование) общения исследователя и опрашиваемого путем регистрации ответов на вопросы, вытекающие из целей и задач исследования. Особенности опроса:

1.Информация, получаемая в ходе опроса отражает реальность, так как она отражается в сознании опрошенных.

2.Содержание информации зависит от установок, настроений, информированности, компетентности опрошенных

3. Данные опроса зависят от содержания вопросов, от организации самого

**Вопросы и задания для контроля знаний по теме:**

1. Что такое «социологическое исследование»? О каких видах социологических исследований Вы знаете?
2. Из каких разделов состоит программа социологического исследования?
3. Охарактеризуйте выборочные метод в социологии
4. Какими бывают методы сбора социологической информации?
5. Какие Вы знаете достоинства и недостатки метода опроса в социологии?
6. Расскажите об основных способах построения анкеты
7. Из каких частей состоит анкета?

**Тема 2. Измерение в социологии**

**Аннотация.** В данной теме студенты знакомятся о понятии измерения в социологии, основных видах шкалирования.

**Ключевые слова:** *измерение, шкалы*

**Методические рекомендации.** Эта тема представляет интерес в том плане, что без понимания основных принципов измерения не возможно правильно провести процедуры анализа данных в социологии.

**Информационные ресурсы.**

1. Татарова Г.Г. Методология анализа данных в социологии. – М.: Изд. дом «Стратегия», 1998. – 222 с.
2. Толстова Ю.Н. Анализ социологических данных: Методология, дескриптивная статистика, изучение связей между номинальными признаками. – М.: Науч. мир, 2000. – 350 с.

**Глоссарий**

***Измерение*** - процедура, с помощью которой объекты измерения, рассматриваемые как носители соответствующих отношений, отображаются в некоторую математическую систему с соответствующими отношениями между элементами

***Шкала*** - алгоритм, с помощью которого устанавливается соответствие между свойствами социального объекта и свойствами сопоставляемых чисел

**Вопросы по теме:**

1. **Система социальных показателей**
2. **Типы шкал в социологии**
3. **Понятия, описывающие качество измерения в социологии**

Под социальным показателем понимается передатчик социальной информации, который находится между исследователем и социальной реальностью. Индикаторы - это доступные наблюдению и измерению характеристики изучаемого социального объекта. Показатели подразделяются на знаково-символические средства, средства наблюдения и измерения, социальные явления и процессы.

 Под измерением в социологическом исследовании понимается процедура, с помощью которой объекты измерения, рассматриваемые как носители соответствующих отношений, отображаются в некоторую математическую систему с соответствующими отношениями между элементами. Это процедура приписывания чисел значениям изучаемого признака, целью которой является получение числовой модели исследования.

Под шкалой понимается тот алгоритм, с помощью которого устанавливается соответствие между свойствами социального объекта и свойствами сопоставляемых чисел.

В социологии обычно выделяют следующие виды шкал:

 Номинальная шкала это шкала соответствия социальных свойств равенства и неравенства соответствующим числовым отношениям с целью отличия одного социального объекта от другого, то есть выявление непересекающихся классов - каждый класс соответствует определенной позиции на шкале.

Порядковая (ранговая) шкала это шкала соответствия социальных свойства равенства или неравенства и их степени выраженности с соответствующими числовыми отношениями. Например, удовлетворенность чем-либо: полностью - 5, скорее да чем нет - 4, трудно сказать - 3, скорее нет - 2, полностью не удовлетворен - 1.

Метрическая шкала это шкала соответствия социальных свойств равенства или неравенства и пропорций внутри них с соответствующими числовыми отношениями. Например, равных пропорций - определение возрастных групп: 16-20 лет - 1, 21-25 лет - 2, 26-30 лет - 3; неравных пропорций - срок службы: 1-2 года - 1, 2-5 лет - 2, 5-20 лет - 3, 20-25 лет - 4, более 25 лет - 5. Если на шкале ноль условный, то это интервальная шкала, если естественный - шкала отношений (возраст, заработанная плата и т.д.).

**Вопросы и задания для контроля знаний по теме:**

1. Что такое измерение в социологии?
2. Какие бывают шкалы в социологии?
3. От чего зависит качество измерения в социологии?

**Тема 3. Этапы анализа данных**

**Аннотация.** В этой теме студенты знакомятся с основными этапами анализа данных в социологии.

**Ключевые слова***:**простая группировка, перекрестная группировка, корреляция, кластерный анализ*

**Методические рекомендации.** Тема очень значима в плане структурирования основных этапов анализа социологической информации, изучения ключевых процедур анализа социологической информации

**Информационные ресурсы.**

1. Татарова Г.Г. Методология анализа данных в социологии. – М.: Изд. дом «Стратегия», 1998. – 222 с.
2. Толстова Ю.Н. Анализ социологических данных: Методология, дескриптивная статистика, изучение связей между номинальными признаками. – М.: Науч. мир, 2000. – 350 с.

**Глоссарий**

***Простая группировка*** — это классификация или упорядочение данных по одному признаку.

***Перекрестная группировка*** — это связывание данных предварительно упорядоченных по двум признакам

***Корреляция*** — **это** статистическая взаимосвязь двух или нескольких случайных величин

***Кластерный анализ*** – это процедура, позволяющая классифицировать различные объекты.

***Вопросы по теме:***

1. Стадии анализа социологической информации
2. Виды статистической обработки социологических данных
3. **Корреляция в социологии**
4. Кластерный анализ.

Стадии анализа социологической информации: 1.общий контроль качества полученной информации («чистка массива»), 2.описание данных, 3. статистическая обработка данных

Группировка— элементарные процедуры упорядочения данных.

Простая группировка (одномерное распределение признаков) — это классификация или упорядочение данных по одному признаку. Группировка решает задачу:

* сжатия исходной информации, ее компактного представления.
* нахождение типичного значения. Вместо большого числа отдельных показателей нам требуется одно значение, которое было бы типичным (репрезентативным) для всей совокупности объектов.

Перекрестная группировка (двумерное распределение признаков) — это связывание данных предварительно упорядоченных по двум признакам

Перекрестная группировка производится с целью:

* обнаружить какие-то взаимозависимости;
* осуществить взаимоконтроль показателей (например, ответов на основной и контрольный вопросы)
* сформировать новый составной показатель (индекс) на основе совмещения двух свойств или состояний объекта, определить направление связей влияния одного явления на другое.

**Корреляция** — **это** статистическая взаимосвязь двух или нескольких случайных величин. Коэффициенты корреляции:

Критерии количественной оценки зависимости между переменными называются *коэффициентами корреляции или мерами связанности.* Значение коэффициента служит показателем интенсивности связи.

Положительную корреляцию следует интерпретировать следующим образом: если значения одной переменной возрастают, то значения другой имеют тенденцию к возрастанию. Чем коэффициент корреляции ближе к 1, тем сильнее эта тенденция, и, наоборот, с приближением коэффициента корреляции к 0 тенденция ослабевает.

Кластерный анализ (от англ. cluster – группа, пучок) – это процедура, позволяющая классифицировать различные объекты. С его помощью можно разбить респондентов на группы, сходные по ряду признаков.

Цель кластерного анализа — классификация объектов на относительно однородные группы исходя из рассматриваемого набора переменных. Объекты в группе относительно схожи между собой и отличаются от объектов в других группах.

**Вопросы и задания для контроля знаний по теме:**

1. Из каких стадий состоит анализ данных в социологии?
2. Чем отличается простая группировка от перекрестной?
3. Что такое корреляция? Какие вы знаете коэффициенты корреляции?
4. Что такое кластерный анализ, где он используется?

**Тема 4. Программное обеспечение в структуре социологического исследования**

**Аннотация.** Данная лекция является важной в понимании студентами программных продуктов, которые существуют на рынке в области анализа и обработки социологической информации.

**Ключевые слова**: MS Excel, SPSS, STATA, STATISTICA

**Методические рекомендации.** Умение разбираться в современных программных продуктов по анализу и обработке социологической информации – целое искусство, так как разные программы предусматривают разные «пакеты» для анализа данных, которые зачастую сложные для понимания студентами гуманитарных факультетов.

**Информационные ресурсы**

1. Бююль А. SPSS: искусство обработки информации. Анализ статистических данных и восстановление закономерностей. – СПб. и др.: ДиаСофт, 2005. – 602 с.
2. Паниотто В.И., Максименко В.С. Количественные методы в социологических исследованиях. – Киев: Наук. думка, 1982. – 272 с.

**Вопросы по теме:**

1. Программное обеспечение в структуре социологического исследования

Использование статистических программ в компьютерной обработке на несколько порядков ускоряет обработку материала и предоставляет в распоряжение исследователя такие методы анализа, которые в ручной обработке не могут быть реализованы. Однако в полной мере эти преимущества могут быть использованы, если исследователь имеет необходимый уровень подготовки в этой области. Обычно, чем мощнее компьютерная программ (чем более широкие у неё возможности), тем больше времени она требует для освоения. Таким образом, затрачивать время на её изучение при редких обращениях к мощному статистическому аппарату не совсем эффективно. Очень часто использование таких программ для решения даже несложных задач также требует определённой суммы умений.

Все программы статистической обработки данных можно разделить на профессиональные, полупрофессиональные (популярные) и специализированные.

**MS Excel** . Самой часто упоминаемой (и используемой) в отечественных статьях является приложение MS Excel из пакета офисных программ компании Microsoft – MS Office. Причины этого кроются в широком распространении этого программного обеспечения, наличии русскоязычной версии, тесной интеграцией с MS Word и PowerPoint. Однако, MS Excel - это электронная таблица с достаточно мощными математическими возможностями, где некоторые статистические функции являются просто дополнительными встроенными формулами. Расчеты сделанные при ее помощи не признаются авторитетными социологическими журналами. Также в MS Excel невозможно построить качественные научные графики.

**SPSS** (Statistical Package for Social Science). Самый часто используемый пакет статистической обработки данных с более чем 30-и летней историей. Отличается гибкостью, мощностью применим для всех видов статистических расчетов применяемых в социологии.

**STATA.** Один из самых популярных в образовательных и научных учреждениях США наряду с SPSS. Официальный сайт. Программа хорошо документирована, издается специальный журнал для пользователей системы. Однако возможности предварительного ознакомления с демо-версией нет.

**STATISTICA** включает большое количество методов статистического анализа (более 250 встроенных функций). Несложный в освоении этот статистический пакет может быть рекомендован для социологических исследований любой сложности.

**NCSS** Программа развивается с 1981 года и рассчитана на непрофессионалов в области статистической обработки. Интерфейс системы многооконный и как следствие этого явления - немного непривычный в использовании. Все действия пользователя сопровождаются подсказками.

**STATGRAPHICS PLUS** . Довольно мощная статистическая программа. Содержит более 250 статистических функций, генерирует понятные, настраиваемые отчеты.

**Вопросы и задания для контроля знаний по теме:**

1. Какие вы знаете программы для обработки и анализа социологической информации?
2. Как программы являются наиболее популярными?

**Практическая работа №1. Обработка социологической информации с помощью Еxсel**

**Аннотация.** Работа в Еxсel – является важным навыком современного социолога, поскольку данная программа является наиболее распространённой и легкой для обработки и анализа социологической информации.

**Ключевые слова:** *Еxсel, база данных, диаграмма*

**Методические рекомендации.**

* Выполнение практических заданий обязательно для каждого студента
* Все практические работы выполняются в компьютерном классе, где инсталлирована лицензионная программа SPSS
* Максимальная оценка за каждую выполненную работу - 5 баллов

**Информационные ресурсы.**

1. Бююль А. SPSS: искусство обработки информации. Анализ статистических данных и восстановление закономерностей. – СПб. и др.: ДиаСофт, 2005. – 602 с.
2. Паниотто В.И., Максименко В.С. Количественные методы в социологических исследованиях. – Киев: Наук. думка, 1982. – 272 с.
3. Татарова Г.Г. Методология анализа данных в социологии. – М.: Изд. дом «Стратегия», 1998. – 222 с.
4. Толстова Ю.Н. Анализ социологических данных: Методология, дескриптивная статистика, изучение связей между номинальными признаками. – М.: Науч. мир, 2000. – 350 с.

**Глоссарий**

***База данных*** – представленная в [объективной](https://ru.wikipedia.org/wiki/%D0%9E%D0%B1%D1%8A%D0%B5%D0%BA%D1%82_(%D1%84%D0%B8%D0%BB%D0%BE%D1%81%D0%BE%D1%84%D0%B8%D1%8F)) форме [совокупность](https://ru.wikipedia.org/wiki/%D0%9C%D0%BD%D0%BE%D0%B6%D0%B5%D1%81%D1%82%D0%B2%D0%BE) самостоятельных материалов (статей, [расчётов](https://ru.wikipedia.org/wiki/%D0%92%D1%8B%D1%87%D0%B8%D1%81%D0%BB%D0%B5%D0%BD%D0%B8%D0%B5), [нормативных актов](https://ru.wikipedia.org/wiki/%D0%9D%D0%BE%D1%80%D0%BC%D0%B0%D1%82%D0%B8%D0%B2%D0%BD%D1%8B%D0%B9_%D0%BF%D1%80%D0%B0%D0%B2%D0%BE%D0%B2%D0%BE%D0%B9_%D0%B0%D0%BA%D1%82), [судебных решений](https://ru.wikipedia.org/wiki/%D0%A1%D1%83%D0%B4%D0%B5%D0%B1%D0%BD%D0%BE%D0%B5_%D1%80%D0%B5%D1%88%D0%B5%D0%BD%D0%B8%D0%B5) и иных подобных материалов), [систематизированных](https://ru.wikipedia.org/wiki/%D0%A1%D0%B8%D1%81%D1%82%D0%B5%D0%BC%D0%B0%D1%82%D0%B8%D0%B7%D0%B0%D1%86%D0%B8%D1%8F) таким образом, чтобы эти материалы могли быть [найдены](https://ru.wikipedia.org/wiki/%D0%9F%D0%BE%D0%B8%D1%81%D0%BA_%D0%B4%D0%B0%D0%BD%D0%BD%D1%8B%D1%85) и обработаны с помощью [электронной вычислительной машины (ЭВМ)](https://ru.wikipedia.org/wiki/%D0%9A%D0%BE%D0%BC%D0%BF%D1%8C%D1%8E%D1%82%D0%B5%D1%80)

***Диаграмма*** -  графическое представление данных линейными отрезками или геометрическими фигурами, позволяющее быстро оценить соотношение нескольких величин. Представляет собой геометрическое [символьное](https://ru.wikipedia.org/wiki/%D0%A1%D0%B8%D0%BC%D0%B2%D0%BE%D0%BB) изображение [информации](https://ru.wikipedia.org/wiki/%D0%98%D0%BD%D1%84%D0%BE%D1%80%D0%BC%D0%B0%D1%86%D0%B8%D1%8F) с применением различных приёмов техники [визуализации](https://ru.wikipedia.org/wiki/%D0%A0%D0%B5%D0%BD%D0%B4%D0%B5%D1%80%D0%B8%D0%BD%D0%B3).

**Вопросы по теме:**

1. Создание базы данных
2. «Чистка массива»
3. Анализ данных
4. Построение диаграммы

1.Создание базы данных – перенос значений из анкеты в базу Excel

2. «Чистка массива» - работа с фильтрами

Фильтрация — это скрытие ненужных данных и показ нужных.

Фильтры могут быть использованы только для одного списка на листе.

1. Укажите ячейки в фильтруемом списке.

2. Меню Данные — Фильтр — Автофильтр.

3. Чтобы отфильтровать строки, содержащие определенное значение, нажмите кнопку со стрелкой в столбце, в котором содержатся искомые данные.

4. Выберите значение в списке.

5. Повторите шаги 3 и 4 для введения дополнительных ограничений значений в других столбцах.

Если данные уже отфильтрованы по одному из столбцов, то при использовании фильтра для другого столбца будут предложены только те значения, которые видны в отфильтрованном списке.

3. Анализ данных

1. В меню Сервис выберите команду Анализ данных.

2. В списке Инструменты анализа выберите нужную строку.

3. Введите входной и выходной диапазоны, затем выберите необходимые параметры.

4. Создание диаграммы

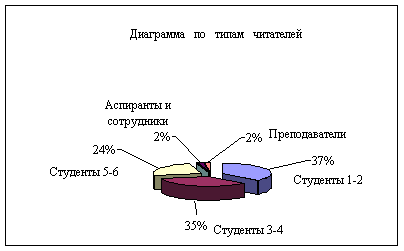
1. Выделите ячейки, содержащие данные, которые должны быть отражены на диаграмме.

2. Если необходимо, чтобы в диаграмме были отражены и названия строк или столбцов, выделите также содержащие их ячейки.

3. Меню Вставка — Диаграмма.

4. Следуйте инструкциям Мастера.

Если на листе присутствуют многоуровневые названия строк или столбцов, то их также можно отразить на диаграмме. При создании диаграммы выделите названия строк и столбцов для каждого уровня. Чтобы при добавлении данных в диаграмму была сохранена заданная иерархия, измените диапазон ячеек, использованных при создании диаграммы.



**Вопросы и задания для контроля знаний по теме:**

На основе проведенного опроса (n=50), студенты сдают в электронном виде базу данных и ее анализ в программе excel. Для этого студенты выполняют следующие действия:

1. Создание базы данных в excel: перенос значений из анкеты в базу Excel

2. «Чистка массива»: работа с фильтрами

3. Анализ данных: вычисление абсолютных значений и процентов по каждой шкале, группировка данных в таблицы

|  |  |  |
| --- | --- | --- |
| Признак | n | % (доля) |
|  |  |  |
|  |  |  |
|  |  |  |

4. Построение диаграмм распределения признаков для каждого вопроса

**Практическая работа №2. Знакомство с пользовательским интерфейсом программы SPSS**

**Аннотация.** На данном практическом занятии **с**туденты знакомятся с ключевой программой по анализу и обработке социологической информации – SPSS.

**Ключевые слова:** *SPSS, база данных*

**Методические рекомендации.**

* Выполнение практических заданий обязательно для каждого студента
* Все практические работы выполняются в компьютерном классе, где инсталлирована лицензионная программа SPSS
* Максимальная оценка за каждую выполненную работу - 5 баллов

**Информационные ресурсы.**

1. Бююль А. SPSS: искусство обработки информации. Анализ статистических данных и восстановление закономерностей. – СПб. и др.: ДиаСофт, 2005. – 602 с.
2. Паниотто В.И., Максименко В.С. Количественные методы в социологических исследованиях. – Киев: Наук. думка, 1982. – 272 с.
3. Татарова Г.Г. Методология анализа данных в социологии. – М.: Изд. дом «Стратегия», 1998. – 222 с.
4. Толстова Ю.Н. Анализ социологических данных: Методология, дескриптивная статистика, изучение связей между номинальными признаками. – М.: Науч. мир, 2000. – 350 с.

**Глоссарий**

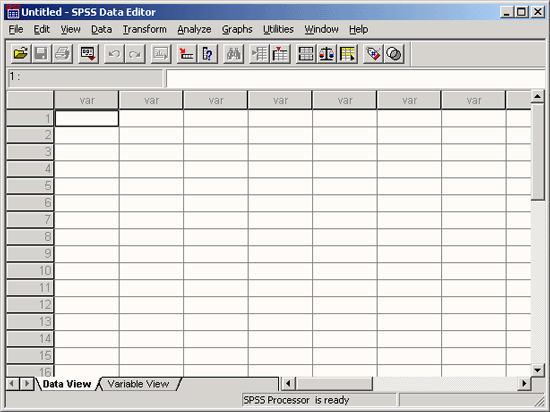
**SPSS Statistics**  — [компьютерная программа](https://ru.wikipedia.org/wiki/%D0%9A%D0%BE%D0%BC%D0%BF%D1%8C%D1%8E%D1%82%D0%B5%D1%80%D0%BD%D0%B0%D1%8F_%D0%BF%D1%80%D0%BE%D0%B3%D1%80%D0%B0%D0%BC%D0%BC%D0%B0) для [статистической обработки](https://ru.wikipedia.org/wiki/%D0%A1%D1%82%D0%B0%D1%82%D0%B8%D1%81%D1%82%D0%B8%D1%87%D0%B5%D1%81%D0%BA%D0%B8%D0%B9_%D0%B0%D0%BD%D0%B0%D0%BB%D0%B8%D0%B7) данных, один из лидеров рынка в области коммерческих статистических продуктов, предназначенных для проведения прикладных исследований в [социальных науках](https://ru.wikipedia.org/wiki/%D0%A1%D0%BE%D1%86%D0%B8%D0%B0%D0%BB%D1%8C%D0%BD%D1%8B%D0%B5_%D0%BD%D0%B0%D1%83%D0%BA%D0%B8).

**Вопросы по теме:**

1. Запуск SPSS. Окна программы
2. Экспорт данных из Excel в SPSS
3. Создание макета данных, просмотр переменных

1. Запуск SPSS. Окна программы

 Запустите SPSS для Windows, дважды щелкнув левой кнопкой мыши на значке SPSS.

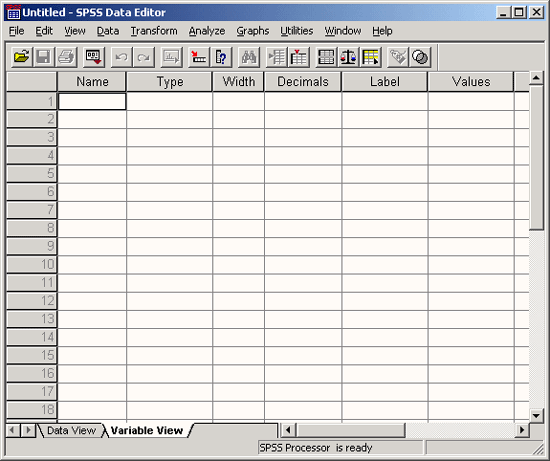


Например, при обработке данных опроса одна строка содержит данные одного респондента. Отдельные столбцы соответствуют отдельным переменным.

Строка меню содержит команды для выполнения почти всех операций, предусмотренных в программе SPSS. Как правило, выполнение команды начинается с появления диалогового окна, в котором пользователю полагается установить значения параметров.

2. Экспорт данных из Excel в SPSS (перенос данных). Перенести только первый вопрос

3. Создание макета данных, просмотр переменных



Вкладка Variable View (Просмотр переменных) позволяет задать структуру файла данных (создать макет данных), то есть определить имена, метки и структуры переменных. Заголовки столбцов представляют собой параметры каждой из переменных: Name (Имя), Type (Тип), Width (Ширина), Decimals (Дробная часть), Label (Метка), Values (Значения), Missing (Пропуски), Columns (Столбцы), Align (Выравнивание), Measure (Измерение).

Имя переменной (Name) - вопрос 1

Метка переменной (Label) - название 1 вопроса (пишем словами)

Метки значений (Values) - варианты ответа на 1 вопрос

Метки значений определяются следующим образом:

* Вначале введите в поле Value (Значение) число "1". Нажмите клавишу <Таb>.
* Введите в поле Value label (Метка значения) текст "мужской".
* Щелкните на кнопке Add (Добавить). Метка значения будет добавлена в список. Для этой цели можно также нажать комбинацию клавиш <Alt>+<h>.
* Повторите эти действия для значений "2" — "женский"

Максимально допустимая длина метки значения составляет 60 знаков.

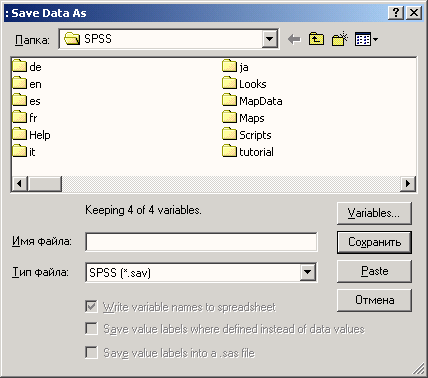
* Подтвердите введенные данные кнопкой ОК, а затем — клавишей <ТаЬ>.

Пропущенные значения: ставим везде «ноль»

Сохранение файла данных

Для того, что бы сохранить созданный файл данных поступите следующим образом:

* Выберите в меню команды File (Файл) Save as... (Сохранить как...) Откроется диалоговое окно Save Data as (Сохранить данные как).

По умолчанию SPSS сохраняет файл данных в текущем каталоге с расширением .sav.

* Задайте имя файла, соответствующее соглашению об именах в DOS. Для рассматриваемого примера мы предлагаем имя файла "opros.sav". Расширение .sav SPSS присваивает файлам данных по умолчанию. Поэтому расширение .sav вводить необязательно.

**Вопросы и задания для контроля знаний по теме:**

Студенты сдают в электронном виде базу данных в SPSS.

Для этого студенты выполняют следующие действия:

1. Создание базы данных в SPSS

2. Экспорт данных из Excel в SPSS

**Практическая работа №3. Одномерные частотные распределения в SPSS**

**Аннотация.** На данном практическом занятиистуденты знакомятся с тем, как проводить одномерный анализ в программе SPSS.

**Ключевые слова:** *среднее, медиана, мода, сумма*

**Методические рекомендации.**

* Выполнение практических заданий обязательно для каждого студента
* Все практические работы выполняются в компьютерном классе, где инсталлирована лицензионная программа SPSS
* Максимальная оценка за каждую выполненную работу - 5 баллов

**Информационные ресурсы.**

1. Бююль А. SPSS: искусство обработки информации. Анализ статистических данных и восстановление закономерностей. – СПб. и др.: ДиаСофт, 2005. – 602 с.
2. Паниотто В.И., Максименко В.С. Количественные методы в социологических исследованиях. – Киев: Наук. думка, 1982. – 272 с.
3. Татарова Г.Г. Методология анализа данных в социологии. – М.: Изд. дом «Стратегия», 1998. – 222 с.
4. Толстова Ю.Н. Анализ социологических данных: Методология, дескриптивная статистика, изучение связей между номинальными признаками. – М.: Науч. мир, 2000. – 350 с.

**Глоссарий**

**Среднее** - мера центральной тенденции

**Медиана -** мера центральной тенденции, не чувствительная к выбросам. Значение, выше и ниже которого попадает половина наблюдений, 50-й процентиль. Если число наблюдений четно, медиана есть арифметическое среднее двух находящихся в середине значений, если выборку упорядочить по убыванию или по возрастанию.

**Мода** - чаще всего встречающееся значение. Если имеется несколько таких значений, каждое из них является модой.

**Сумма** - сумма значений для всех наблюдений с непропущенными значениями.

**Вопросы по теме:**

1. Одномерный анализ данных в SPSS

|  |  |
| --- | --- |
|  | Выберите в меню: Анализ > Описательные статистики > Частоты…  Выберите одну или несколько категориальных или количественных переменных.  Дополнительно Вы можете:   * Щелкнуть мышью по кнопке Статистики, чтобы задать вычисление описательных статистик для количественных переменных. * Щелкнуть мышью по кнопке Диаграммы, чтобы задать вывод столбиковых диаграмм, круговых диаграмм и гистограмм.   Щелкнуть мышью по кнопке Формат, чтобы задать порядок, в котором будут выводиться результаты.  Выберите в меню: Анализ > Описательные статистики > Частоты…   * В диалоговом окне Частоты щелкните по кнопке Диаграммы |

* Стандартное отклонение. Мера разброса вокруг среднего, выраженная в тех же единицах измерения, что и наблюдения. Равна корню квадратному из дисперсии. При нормальном распределении 68% наблюдений укладывается в одно стандартное отклонение от среднего, а 95% процентов в 2 стандартных отклонения. Если, например, средний возраст равен 45 годам со стандартным отклонением 10, то 95% наблюдений должны оказаться между 25 и 65 годами при нормальном распределении.
* Дисперсия. Мера рассеяния относительно среднего. Равна сумме квадратов отклонений от среднего, деленной на число, на единицу меньшее числа наблюдений. Единица измерения - квадрат единицы измерения переменной.
* Размах. Разность между наибольшим и наименьшим значениями числовой переменной; максимум минус минимум.
* Минимум. Наименьшее значение числовой переменной.
* Максимум. Наибольшее значение числовой переменной.
* Стандартная ошибка среднего. Мера того, насколько сильно среднее может различаться у разных выборок, взятых из одного распределения. Можно применять для грубого сравнения наблюденного среднего с гипотетическим значением (то есть можно заключить, что два значения различаются, если отношение их разности к стандартному отклонению меньше -2 или больше +2).

**Вопросы и задания для контроля знаний по теме:**

Студенты сдают в электронном виде базу данных в формате Microsoft Word c одномерным анализом данных, диаграммами и интерпретацией полученных данных по каждому вопросу. Для этого студенты выполняют следующие действия:

1. Одномерный анализ данных для каждого вопроса

2. Построение диаграмм

3. Описание данных, сравнение и сопоставление для каждого вопроса

**Практическая работа №4. Двумерные частотные распределения в SPSS**

**Аннотация.** На данном практическом занятии **с**туденты знакомятся с тем, как проводить двумерный анализ в программе SPSS.

**Ключевые слова:** *Коэффициент сопряженности, меры связи фи и V Крамера, лямбда, коэффициент неопределенности, гамма.*

**Методические рекомендации.**

* Выполнение практических заданий обязательно для каждого студента
* Все практические работы выполняются в компьютерном классе, где инсталлирована лицензионная программа SPSS
* Максимальная оценка за каждую выполненную работу - 5 баллов

**Информационные ресурсы.**

1. Бююль А. SPSS: искусство обработки информации. Анализ статистических данных и восстановление закономерностей. – СПб. и др.: ДиаСофт, 2005. – 602 с.
2. Паниотто В.И., Максименко В.С. Количественные методы в социологических исследованиях. – Киев: Наук. думка, 1982. – 272 с.
3. Татарова Г.Г. Методология анализа данных в социологии. – М.: Изд. дом «Стратегия», 1998. – 222 с.
4. Толстова Ю.Н. Анализ социологических данных: Методология, дескриптивная статистика, изучение связей между номинальными признаками. – М.: Науч. мир, 2000. – 350 с.

**Глоссарий**

**Коэффициент сопряженности**. Мера ассоциативной связи, основанная на хи-квадрат. Значение этого коэффициента заключено между 0 и 1, причем 0 означает отсутствие связи между переменными строки и столбца, а значение, близкое к 1, - высокую степень связи между этими переменными. Максимально возможное значение зависит от числа строк и столбцов в таблице.

**Меры связи фи и V Крамера.** Фи - это мера связи, равная корню квадратному из значения хи-квадрат, деленного на объем выборки. V Крамера - это мера связи, основанная на статистике хи-квадрат.

**Лямбда.** Мера ассоциации, которая отражает относительное снижение ошибки, когда значения независимой переменной используются для предсказания значений зависимой переменной. Значение 1 означает, что независимая переменная полностью предсказывает значения зависимой. Значение 0 означает, что независимая переменная абсолютно бесполезна для предсказания зависимой.

**Коэффициент неопределенности.** Мера связи, указывающая относительное снижение ошибки в случае, когда значения одной переменной используются для предсказания значений другой. Например, значение 0.83 указывает на то, что знание одной переменной уменьшает ошибку в предсказании значений другой на 83%. Вычисляются как симметричная, так и несимметричная версии коэффициента неопределенности.

**Гамма**. Симметричная мера связи между двумя порядковыми переменными, значения которой меняются между -1 и 1. Значения, близкие по абсолютной величине к 1, указывают на сильную связь переменных, значения, близкие к 0, - на слабую связь или ее отсутствие. Для таблиц сопряженности двух переменных вычисляется гамма нулевого порядка. Если же таблица сопряженности включает более двух переменных, для каждой подтаблицы вычисляется условная гамма.

**Вопросы по теме:**

1.Двумерный анализ данных в SPSS

Чтобы помочь вам выявить структуры в данных, которые могут повлиять на результаты критерия хи-квадрат, процедура Таблицы сопряженности выводит ожидаемые значения частот и три типа остатков (отклонений), которые выступают как меры различия между ожидаемыми и наблюденными частотами. Каждая ячейка таблицы может содержать любую комбинацию выбранных частот, процентов и остатков.

Отметьте Хи-квадрат, чтобы получить значения критериев хи-квадрат Пирсона, хи-квадрат отношения правдоподобия, точного критерия Фишера и критерия хи-квадрат с поправкой Йетса (с поправкой на непрерывность) для таблиц, образованных двумя строками и двумя столбцами. Для таблиц 2 × 2 точный критерий Фишера вычисляется в том случае, когда таблица, которая не является результатом наличия пропущенных строк или столбцов в таблице большего размера, имеет ожидаемое значение меньше 5 хотя бы в одной ячейке. Для всех остальных таблиц размерности 2 × 2 рассчитывается критерий хи-квадрат с поправкой Йетса. Для таблиц с любым числом строк и столбцов отметьте Хи-квадрат, чтобы вывести значения хи-квадрата Пирсона и хи-квадрат отношения правдоподобия. Если обе переменные в таблице являются количественными, то при пометке элемента Хи-квадрат рассчитывается критерий линейно-линейной связи.

Для таблиц с упорядоченными переменными по строкам и столбцам при пометке элемента Корреляции вычисляются значения коэффициента корреляции Спирмана - ро (только для числовых данных). ро Спирмана является мерой связи между порядковыми переменными. Если обе переменные в таблице (факторы) являются числовыми, параметр Корреляции позволяет вычислить коэффициент корреляции Пирсона r, который характеризует силу линейной связи между переменными.

Номинальные

Для номинальных данных (которые не имеют естественного порядка - например, католическое, протестантское, иудейское вероисповедание) можно выбрать одну из следующих статистик: Фи (коэффициент) и V Крамéра, Коэффициент сопряженности,Лямбда (симметричное и асимметричное значения лямбда, статистика тау Гудмана и Краскала), Коэффициент неопределенности.

Порядковые

Для таблиц, в которых как строки, так и столбцы содержат упорядоченные значения, пометьте Гамма (нулевого порядка для двумерных таблиц и условное для таблиц размерности от 2 до 10), тау-b Кендалла и тау-c Кендалла. Для предсказания категорий столбца по категориям строки, пометьте d Сомерса.

Номининально-интервальные

В ситуации, когда одна из переменных номинальная, а другая - количественная, выберите статистику Эта. Значения категориальной переменной должны быть закодированы числами.

**Вопросы и задания для контроля знаний по теме:**

Студенты сдают в электронном виде базу данных в формате Microsoft Word c двумерным анализом данных и интерпретацией полученных данных по каждому вопросу. Для этого студенты выполняют следующие действия:

1. Двумерный анализ данных для каждого вопроса

2. Описание данных, сравнение и сопоставление для каждого вопроса

**Практическая работа №3. Одномерные частотные распределения в SPSS**

**Аннотация.** На данном практическом занятиистуденты знакомятся с тем, как проводить одномерный анализ в программе SPSS.

**Ключевые слова:** *среднее, медиана, мода, сумма*

**Методические рекомендации.**

* Выполнение практических заданий обязательно для каждого студента
* Все практические работы выполняются в компьютерном классе, где инсталлирована лицензионная программа SPSS
* Максимальная оценка за каждую выполненную работу - 5 баллов

**Информационные ресурсы.**

1. Бююль А. SPSS: искусство обработки информации. Анализ статистических данных и восстановление закономерностей. – СПб. и др.: ДиаСофт, 2005. – 602 с.
2. Паниотто В.И., Максименко В.С. Количественные методы в социологических исследованиях. – Киев: Наук. думка, 1982. – 272 с.
3. Татарова Г.Г. Методология анализа данных в социологии. – М.: Изд. дом «Стратегия», 1998. – 222 с.
4. Толстова Ю.Н. Анализ социологических данных: Методология, дескриптивная статистика, изучение связей между номинальными признаками. – М.: Науч. мир, 2000. – 350 с.

**Глоссарий**

**Среднее** - мера центральной тенденции

**Медиана -** мера центральной тенденции, не чувствительная к выбросам. Значение, выше и ниже которого попадает половина наблюдений, 50-й процентиль. Если число наблюдений четно, медиана есть арифметическое среднее двух находящихся в середине значений, если выборку упорядочить по убыванию или по возрастанию.

**Мода** - чаще всего встречающееся значение. Если имеется несколько таких значений, каждое из них является модой.

**Сумма** - сумма значений для всех наблюдений с непропущенными значениями.

**Вопросы по теме:**

1. Одномерный анализ данных в SPSS

|  |  |
| --- | --- |
|  | Выберите в меню: Анализ > Описательные статистики > Частоты…  Выберите одну или несколько категориальных или количественных переменных.  Дополнительно Вы можете:   * Щелкнуть мышью по кнопке Статистики, чтобы задать вычисление описательных статистик для количественных переменных. * Щелкнуть мышью по кнопке Диаграммы, чтобы задать вывод столбиковых диаграмм, круговых диаграмм и гистограмм.   Щелкнуть мышью по кнопке Формат, чтобы задать порядок, в котором будут выводиться результаты.  Выберите в меню: Анализ > Описательные статистики > Частоты…   * В диалоговом окне Частоты щелкните по кнопке Диаграммы |

* Стандартное отклонение. Мера разброса вокруг среднего, выраженная в тех же единицах измерения, что и наблюдения. Равна корню квадратному из дисперсии. При нормальном распределении 68% наблюдений укладывается в одно стандартное отклонение от среднего, а 95% процентов в 2 стандартных отклонения. Если, например, средний возраст равен 45 годам со стандартным отклонением 10, то 95% наблюдений должны оказаться между 25 и 65 годами при нормальном распределении.
* Дисперсия. Мера рассеяния относительно среднего. Равна сумме квадратов отклонений от среднего, деленной на число, на единицу меньшее числа наблюдений. Единица измерения - квадрат единицы измерения переменной.
* Размах. Разность между наибольшим и наименьшим значениями числовой переменной; максимум минус минимум.
* Минимум. Наименьшее значение числовой переменной.
* Максимум. Наибольшее значение числовой переменной.
* Стандартная ошибка среднего. Мера того, насколько сильно среднее может различаться у разных выборок, взятых из одного распределения. Можно применять для грубого сравнения наблюденного среднего с гипотетическим значением (то есть можно заключить, что два значения различаются, если отношение их разности к стандартному отклонению меньше -2 или больше +2).

**Вопросы и задания для контроля знаний по теме:**

Студенты сдают в электронном виде базу данных в формате Microsoft Word c одномерным анализом данных, диаграммами и интерпретацией полученных данных по каждому вопросу. Для этого студенты выполняют следующие действия:

1. Одномерный анализ данных для каждого вопроса

2. Построение диаграмм

3. Описание данных, сравнение и сопоставление для каждого вопроса

**Практическая работа №5. Работа с социологической информацией в Интернете**

**Аннотация.** На данном практическом занятии **с**туденты учатся навыкам анализа данных в Интернете, а также познакомятся с Единым архивом социологических и экономических данных.

**Ключевые слова:** *Единый архив социологических и экономических данных*

**Методические рекомендации.**

* Выполнение практических заданий обязательно для каждого студента
* Все практические работы выполняются в компьютерном классе, где инсталлирована лицензионная программа SPSS
* Максимальная оценка за каждую выполненную работу - 5 баллов

**Информационные ресурсы.**

1. Бююль А. SPSS: искусство обработки информации. Анализ статистических данных и восстановление закономерностей. – СПб. и др.: ДиаСофт, 2005. – 602 с.
2. Паниотто В.И., Максименко В.С. Количественные методы в социологических исследованиях. – Киев: Наук. думка, 1982. – 272 с.
3. Татарова Г.Г. Методология анализа данных в социологии. – М.: Изд. дом «Стратегия», 1998. – 222 с.
4. Толстова Ю.Н. Анализ социологических данных: Методология, дескриптивная статистика, изучение связей между номинальными признаками. – М.: Науч. мир, 2000. – 350 с.

**Глоссарий**

**Единый архив социологических и экономических данных –** Единый архив экономических и социологических данных (ЕАЭСД) – организация, обеспечивающая свободный и открытый доступ к результатам эмпирических исследований в сфере наук об обществе. Архив был основан в 2000 году. К настоящему моменту в коллекциях Архива накоплено порядка 1000 социологических исследований и более 125 временных рядов основных показателей российской экономики.

**Вопросы по теме:**

* 1. Единый архив социологических и экономических данных
  2. Демонстрация работы в архиве

Единый архив социологических и экономических данных возник на базе двух структур – Единого архива социологических данных (ЕАСД) и Центра статистических данных (ЦСД).

ЕАСД начал свою историю в сентябре 2000 года. Инициатором его создания выступил ВЦИОМ под руководством Юрия Александровича Левады, самую действенную поддержку развитию проекта оказала Татьяна Ивановна Заславская. К проекту присоединились ведущие исследовательские организации, ставшие первыми депозиторами архива: Институт социологии Российской академии наук, Фонд «Общественное мнение», Институт комплексных социологических исследований, РОМИР под руководством Елены Башкировой, Институт экономики и организации промышленного производства Сибирского отделения РАН, безвозмездно передавшие в ЕАСД свои исследования. С 2002 по 2009 год ЕАСД осуществлял свою работу как структурное подразделение Независимого института социальной политики (НИСП). Финансовую поддержку проекту оказали Фонд Форда (с 2000 года) и Фонд Макартуров (с 2006 года). Коллекции социологических данных ЕАСД с первых дней работы были представлены в сети Интернет, что обеспечивало свободный доступ к данным. Сайт поддерживался на русском и английском языке.

Экономическая часть коллекций ЕАЭСД возникла и развивалась в рамках ВШЭ. Идея создания баз данных по российской экономике была предложена Евгением Григорьевичем Ясиным в 1999 году. В следующем году был учрежден Институт информационного развития, занимавшийся, в том числе, поддержкой и сопровождением баз данных по экономике. В 2001 году экономические базы данных стали открытым информационным ресурсом, доступ к которому осуществлялся через Интернет, в 2002 году появилась англоязычная версия системы.

Решение о создании единого научно-исследовательского подразделения было принято Ученым советом ВШЭ в декабре 2009 года. С этого момента начинается новый этап развития Единого архива экономических и социологических данных. Но основные принципы работы – обеспечение сохранности эмпирических данных и организация свободного доступа – остались прежними.

**Вопросы и задания для контроля знаний по теме:**

Студенты выбирают тематическую базу данных для анализа. На основе имеющихся данных необходимо описать методику, которая использовалась в опросе, придумать гипотезу, а также обобщить имеющиеся данные. В файле следует указать фамилии исполнителей.