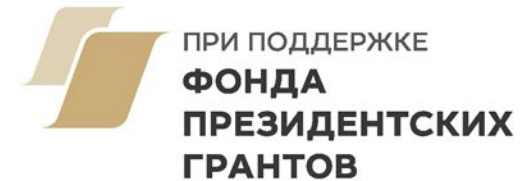




ПРИ ПОДДЕРЖКЕ
ФОНДА
ПРЕЗИДЕНТСКИХ
ГРАНТОВ

Тема лекции:
**«Особенности работы с Excel, Google
Forms и SPSS»**

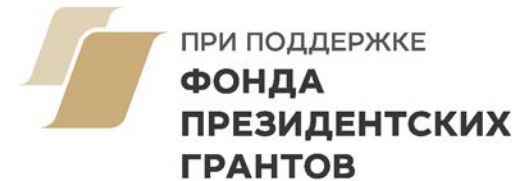
Особенности работы с пакетом SPSS



SPSS — Statistical Package for the Social Science (статистический пакет для социальных наук)

Наряду с другими статистическими пакетами (*Statistica*, *STATA*, *SAS*) широко используется специалистами в сфере исследований (социология, психология, маркетинг, медицина и пр.) для обработки и анализа количественных данных.

Особенности работы с пакетом SPSS



SPSS – это аббревиатура от Statistical Package for the Social Science (статистический пакет для социальных наук).

Как следует из названия, SPSS представляет собой множество различных программ, предназначенных для анализа данных в социальных науках.

Эти программы позволяют организовывать ввод данных, гибко менять их структуру, применять к ним самые современные методы обработки или их последовательность и получать результаты в удобной и наглядной форме.

Все это множество программ объединено в единую систему, обеспечивающую простой и дружественный диалог с исследователем и снабженную исчерпывающей справочной поддержкой. Благодаря такой дружественности система SPSS легко доступна для освоения даже тем, кто имеет минимальные навыки владения компьютером.

Запуск программы, интерфейс, принципы работы

Запуск SPSS Start → All Programs → SPSS Inc → IBM SPSS Statistics

The screenshot displays the IBM SPSS Statistics software interface. At the top, the title bar reads "Без имени1 [Набор данных0] - Редактор данных IBM SPSS Statistics". Below the title bar is a menu bar with options: "Файл", "Правка", "Вид", "Данные", "Преобразовать", "Анализ", "Прямой маркетинг", "Графика", "Сервис", "Окно", "Справка". A toolbar with various icons is located below the menu bar. The main workspace is a data editor window with a grid. The grid has columns labeled: "Имя", "Тип", "Ширина", "Десятич...", "Метка", "Значения", "Пропущенн...", "Ширина ...", "Выравнивание", "Шкала", and "Роль". The rows are numbered from 1 to 36. The right side of the window is a large, empty light blue area. At the bottom of the window, there are tabs for "Данные" and "Переменные". The Windows taskbar is visible at the very bottom, showing the system tray with the date "12.07.2021" and time "11:57".

Запуск программы, интерфейс, принципы работы

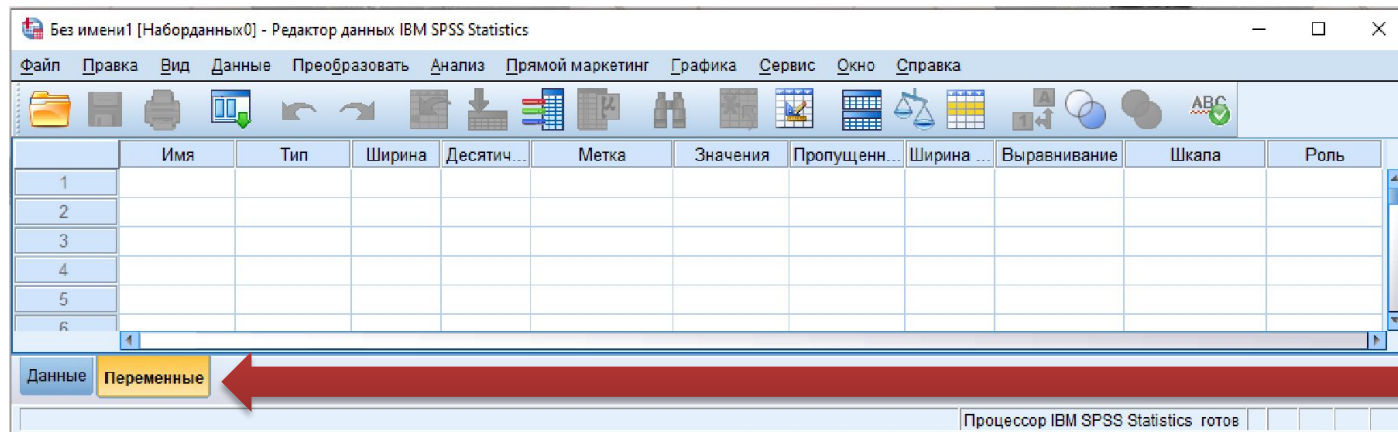
При запуске SPSS пользователю открывается окно для ввода, редактирования и просмотра данных исследования. Данные сохраняются в файле с расширением *.sav

Вкладка «Data view» (Окно данных)

- Окно ввода данных
- Columns: variables (переменные)
- Rows: cases (наблюдения)

Вкладка «Variable view» (Окно переменных)

- Окно создания и настройки переменных
- Variables (список переменных)
- Параметры каждой из переменных



В отличие, например, от Excel, результаты вычислений, анализа данных и их представление в визуальном виде производится в отдельном окне. Окно отображает историю команд (слева), вывод результатов расчетов и возникающие ошибки расчетов (справа). Сохранение осуществляется в файл *.spv или *.spv

The screenshot shows the IBM SPSS Statistics Viewer interface. The left pane displays the command history, and the right pane shows the output of a frequency analysis.

Command History:

```
GET
  FILE='C:\Users\HP_user_5\Desktop\Общая папка\Документы (Альбина)
  \Все документы\Все документы\5\база предприниматели суджанский р-н.sav'.
DATASET NAME Наборданный1 WINDOW=FRONT.
FREQUENCIES VARIABLES=B22
  /ORDER=ANALYSIS.
```

Output:

Частоты

[Наборданный1] C:\Users\HP_user_5\Desktop\Общая папка\Документы (Альбина)\Все документы\Все документы\5\база предпри...

Статистики

В каком населенном пункте Вы проживаете?

N	Валидные	24
	Пропущенные	0

В каком населенном пункте Вы проживаете?

	Частота	Процент	Валидный процент	Накопленный процент
Валидные Суджанский	24	100,0	100,0	100,0

Процессор IBM SPSS Statistics: rotov

Создание файлов данных. Настройка переменных

3 ЭТАПА СОЗДАНИЯ БАЗЫ ДАННЫХ ДЛЯ АНАЛИЗА

До начала ввода данных в SPSS необходимо создать макет (структуру) переменных (на основе, например, анкеты). В этом случае вопросы ложатся в основу переменных. У каждой переменной – свои настройки (имя, допустимые значения, тип шкалы и др.)

Структуру базы данных правильнее определить на этапе планирования исследования и разработки инструментария (например, анкеты) в соответствии с гипотезой и задачами исследования.

Шаг 1.
Задание имён
переменных



Шаг 2.
Определение
их параметров

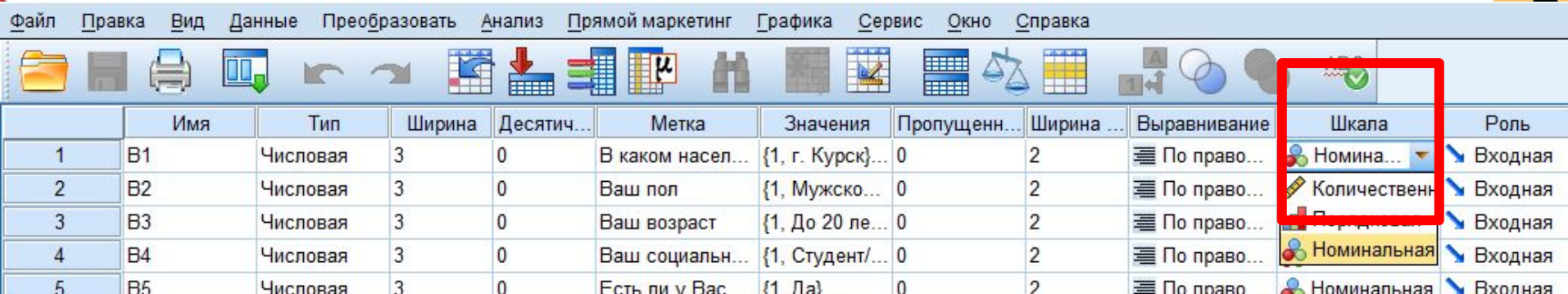


Шаг 3. Ввод
данных

Шкалы

Шкала (Scale) – правило, определяющее, каким образом в процессе измерения каждому изучаемому объекту ставится в соответствие некоторое число или символы.

Шкалирование (Scaling) – процесс создания континуума (последовательного ряда), на котором размещаются измеряемые объекты.



	Имя	Тип	Ширина	Десятич...	Метка	Значения	Пропущенн...	Ширина ...	Выравнивание	Шкала	Роль
1	B1	Числовая	3	0	В каком насел...	{1, г. Курск}...	0	2	≡ По право...	Номина...	Входная
2	B2	Числовая	3	0	Ваш пол	{1, Мужско...	0	2	≡ По право...	Количественн...	Входная
3	B3	Числовая	3	0	Ваш возраст	{1, До 20 ле...	0	2	≡ По право...	Номинальная	Входная
4	B4	Числовая	3	0	Ваш социальн...	{1, Студент/...	0	2	≡ По право...	Номинальная	Входная
5	B5	Числовая	3	0	Есть ли у Вас ...	{1, Да}...	0	2	≡ По право...	Номинальная	Входная

Типы шкал

Номинальная шкала (Nominal) — шкала наименований, которая состоит из значений признаков, не упорядоченных по степени возрастания или убывания. **Пример:** *национальность, профессия, семейное положение, пол и т.д.*

Порядковая шкала (Ordinal) — градации располагаются в определенном порядке относительно возрастания либо убывания интенсивности свойства. **Пример:** *переменная «Курение» со значениями (1 = некурящий; 2 = изредка курящий; 3 = интенсивно курящий; 4 = очень интенсивно курящий). Переменная сортирована в порядке значимости снизу вверх: умеренный курильщик курит больше, нежели некурящий, а сильно курящий — больше, чем умеренный курильщик и т.д., поэтому порядковая шкала.*

Интервальные шкалы (Interval) — основаны на процедурах, обеспечивающих равные или примерно равные расстояния между градациями переменной. В данном случае сравниваются не значения переменных, а расстояния между значениями. **Пример:** *температура, измеренная в градусах Цельсия. Можно не только сказать, что температура 30 градусов выше, чем 20 градусов, но и то, что увеличение температуры с 10 до 30 градусов вдвое больше увеличения температуры от 20 до 30 градусов.*

Шкалы отношений (Метрические) — соответствуют всем требованиям, предъявляемым к шкалам более низких классов. **Пример:** *возраст. Если Макс 30 лет, а Сергею 60, можно сказать, что Сергей вдвое старше Макса.*

Ввод данных

Расчет описательных статистик в SPSS

Описательные статистики (Descriptive Statistics) - это основные статистические параметры, которыми можно описать имеющееся распределение данных, если оно носит характер близкий к нормальному распределению. *Все данные условно делятся на два больших класса:*



Рассмотрим, как вычисляются описательные статистики для двух типов данных.

Описательные статистики для дискретных данных

Сначала загрузите файл job.sav, выбрав команды меню File / Open...((Файл) Открыть...).

Выберите в меню команды (**Анализ**) / (**Описательные статистики**) / (**Частоты**). Появится диалоговое окно Частоты.

Кнопкой со стрелкой перенесите переменную EDUCAT в список выходных переменных и подтвердите операцию кнопкой ОК.

Файл Правка Вид Данные Преобразовать Анализ Прямой маркетинг Графика Сервис Окно Справка

	Имя	Тип	Ширина
1	B1	Числовая	3
2	B2	Числовая	3
3	B3	Числовая	3
4	B4	Числовая	3
5	B5	Числовая	3
6	B6	Числовая	3

Отчеты
Описательные статистики
Таблицы
Сравнение средних
Общая линейная модель
Обобщенные линейные модели
Смешанные модели
Корреляции

Частоты...
Частоты

Переменные:
Ваш пол [B2]

Вывести частотные таблицы

Статистики...
Диagramмы...
Формат
Бутстреп...

ОК Вставка Сброс Отмена Справка

Описательные статистики для дискретных данных

1. Каждая строка частотной таблицы описывает одно возможное значение.
2. Первый столбец содержит метки отдельных значений (уровень образования).
3. Во втором столбце под заголовком «Частота» приведена частота каждого из вариантов уровня образования.

Пример: 38 респондентов имеют «неполное среднее образование», а у большинства опрошенных высшее образование – 764.

Статистики

Ваше образование

N	Валидные	1326
	Пропущенные	0

Ваше образование

		Частота	Процент	Валидный процент	Накопленный процент
Валидные	Неполное среднее	38	2,9	2,9	2,9
	Общее среднее	66	5,0	5,0	7,8
	Среднее специальное	324	24,4	24,4	32,3
	Неполное высшее	134	10,1	10,1	42,4
	Высшее	764	57,6	57,6	100,0
	Итого	1326	100,0	100,0	

Описательные статистики для дискретных данных

4. В третьем столбце показана процентная частота каждого варианта образования.

5. В четвертом столбце дано допустимое процентное значение (исключены потерянные данные).

6. Последний столбец «Накопленный процент» содержит сумму процентных частот.

Статистики

Ваше образование

N	Валидные	1326
	Пропущенные	0

Ваше образование

		Частота	Процент	Валидный процент	Накопленный процент
Валидные	Неполное среднее	38	2,9	2,9	2,9
	Общее среднее	66	5,0	5,0	7,8
	Среднее специальное	324	24,4	24,4	32,3
	Неполное высшее	134	10,1	10,1	42,4
	Высшее	764	57,6	57,6	100,0
	Итого	1326	100,0	100,0	

Анализ множественных ответов

При анализе и кодировании множественных ответов (вопросы, на которые можно дать несколько ответов одновременно) используются два метода:

Метод множественной дихотомии

(для каждой из возможностей ответа определяется отдельная переменная)



Метод множественных категорий

(должно быть известно максимальное количество возможных ответов)

Метод множественных категорий

Пример: В массиве данных job.sav был вопрос: «На что, по Вашему мнению, должна быть в первую очередь направлена работа по развитию конкуренции в Курской области?». **Варианты ответов:** 1. создание условий для того, чтобы увеличения юридических и физических лиц, 2. создание системы информирования населения о работе розничных компаний, защите прав потребителей и состояния конкуренции и т.д.

Кодирование переменных:

Максимальное количество возможных ответов равно 3. Каждая из трех переменных кодируется одинаковыми категориями. Не зависимо от количества данных ответов область этих переменных заполняется слева направо.

Определение наборов в SPSS:

- «Анализ» → «Множественные ответы» → «Определить наборы»
- Выбрать переменные Q6_1 – Q6_5
- Задать категориальную кодировку
- Выбрать диапазон значений: 1:13

Задать наборы множественных ответов

Параметры набора

Оцените УРОВЕ...
Оцените УРОВЕ...
Оцените УДОБС...
VAR00001
VAR00002
VAR00003
VAR00004
VAR00005
VAR00006

Переменные в наборе:
На что, по Вашему ...
На что, по Вашему ...
На что, по Вашему ...

Переменные кодируются как:
 Дихотомии Подсчитываемое значение:
 Категории Диапазон 1 до 13

Имя: B15
Метка: Работа по развитию

Наборы множественных ответов:

Добавить
Изменить
Удалить

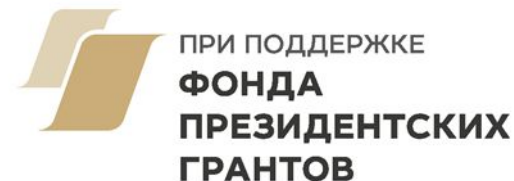
Примечание: Наборы, заданные здесь, доступны только для процедур Частоты для множественных ответов и Таблицы сопряженности для множественных ответов. Для задания наборов множественных ответов, которые можно будет использовать в других процедурах, следует воспользоваться пунктом Задать наборы множественных ответов в меню Данные

Закреть Справка

Дополнительные материалы

- Бююль А., Цеффель П. SPSS: искусство обработки информации. – М., 2005
- Наследов А. IBM SPSS Statistics 20 и AMOS: профессиональный статистический анализ данных. – СПб., 2013
- Измерение в социологии: учеб. пособие / А.П. Кулаков; Новосиб.гос. архитектур.-строит. ун-т. – Новосибирск : НГАСУ (Сибстрин), 2005
- Крыштановский А.О. Анализ социологических данных с помощью пакета SPSS – Москва., 2006
- Е.В. Дорогонько Обработка и анализ социологическх данных с помощью пакета SPSS – Сургут Издательский центр СурГУ., 2010
- Иллюстрированный самоучитель по SPSS [Электронный ресурс]. URL: www.learnspss.ru
- SPSS: обработка статистической информации [Электронный ресурс]. URL: www.ispss.ru
- Иллюстрированный самоучитель по SPSS [Электронный ресурс] / Компьютерная документация Hardline.RU. – М., 2006. – URL : <http://www.hardline.ru>

Выводы:



Специализированные пакеты программного обеспечения, одним из наиболее популярных среди которых является SPSS (PASW), призваны облегчить и ускорить работу, обеспечивает быструю и точную обработку данных.

Основной особенностью этой программы служит то, что результаты анализа можно наглядно представлять в виде таблиц и диаграмм различных типов, распространять сетевым пользователям, внедрять полученные результаты в другие программные системы.