

МИНОБРНАУКИ РОССИИ

**федеральное государственное бюджетное образовательное
учреждение высшего образования**

«ЧЕРЕПОВЕЦКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»



Гуманитарный институт

Кафедра социологии и социальных технологий

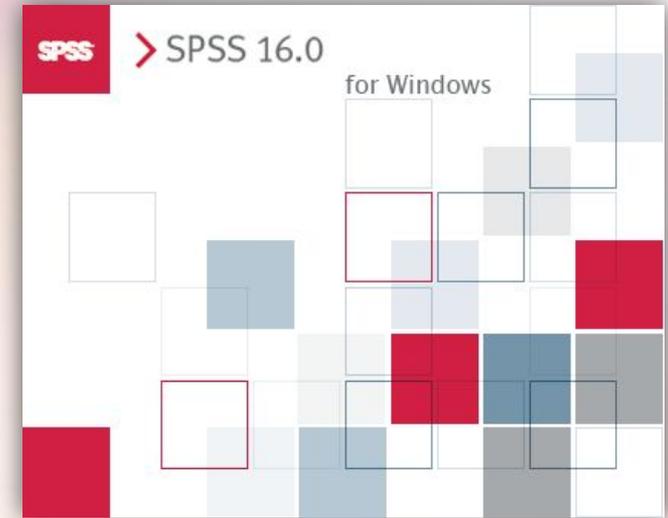
Дисциплина Прикладные компьютерные программы в обработке социологических данных

Лекция 1. Вводная

Череповец
2017 г.

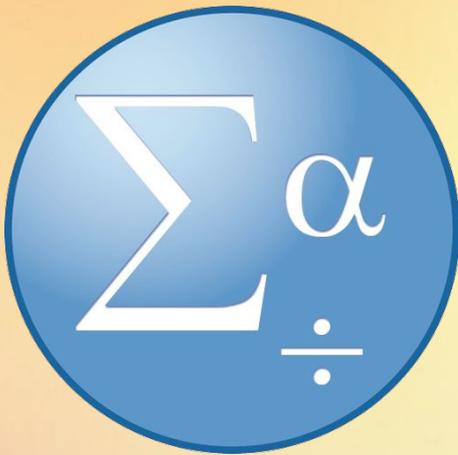
SPSS

SPSS (Statistical Package for the Social Sciences — статистический пакет для социальных наук) — компьютерная программа для статистической обработки данных, предназначенная для проведения прикладных исследований в социальных науках.



Общие функциональные возможности пакета SPSS:

- Оперативное получение аналитической информации – ключевых факторов, взаимосвязей и тенденций в исследуемых данных.
- Наглядное представление результатов в виде таблиц и диаграмм.
- Распространение полученных результатов сетевым пользователям – заказчикам проведенных исследований.
- Поддержка процессов оперативного формирования научно-обоснованных оптимальных управленческих решений.



Основные блоки SPSS

Редактор данных – гибкая система, внешне похожая на электронную таблицу, для определения, ввода, редактирования и просмотра данных.

Средство просмотра – упрощает просмотр результатов, позволяя показывать и скрывать отдельные элементы вывода, изменять порядок вывода результатов, перемещать готовые к презентации таблицы и диаграммы в другие приложения и получать их из приложений.

Многомерные мобильные таблицы – служат для отображения результатов анализа.



Высококачественные графика – средство формирования диаграмм.

Доступ к базам данных – конструктор чтения баз данных, позволяющий загрузить данные из любого источника.

Преобразование данных – средство преобразования данных, помогает готовить данные к анализу.

SPSS

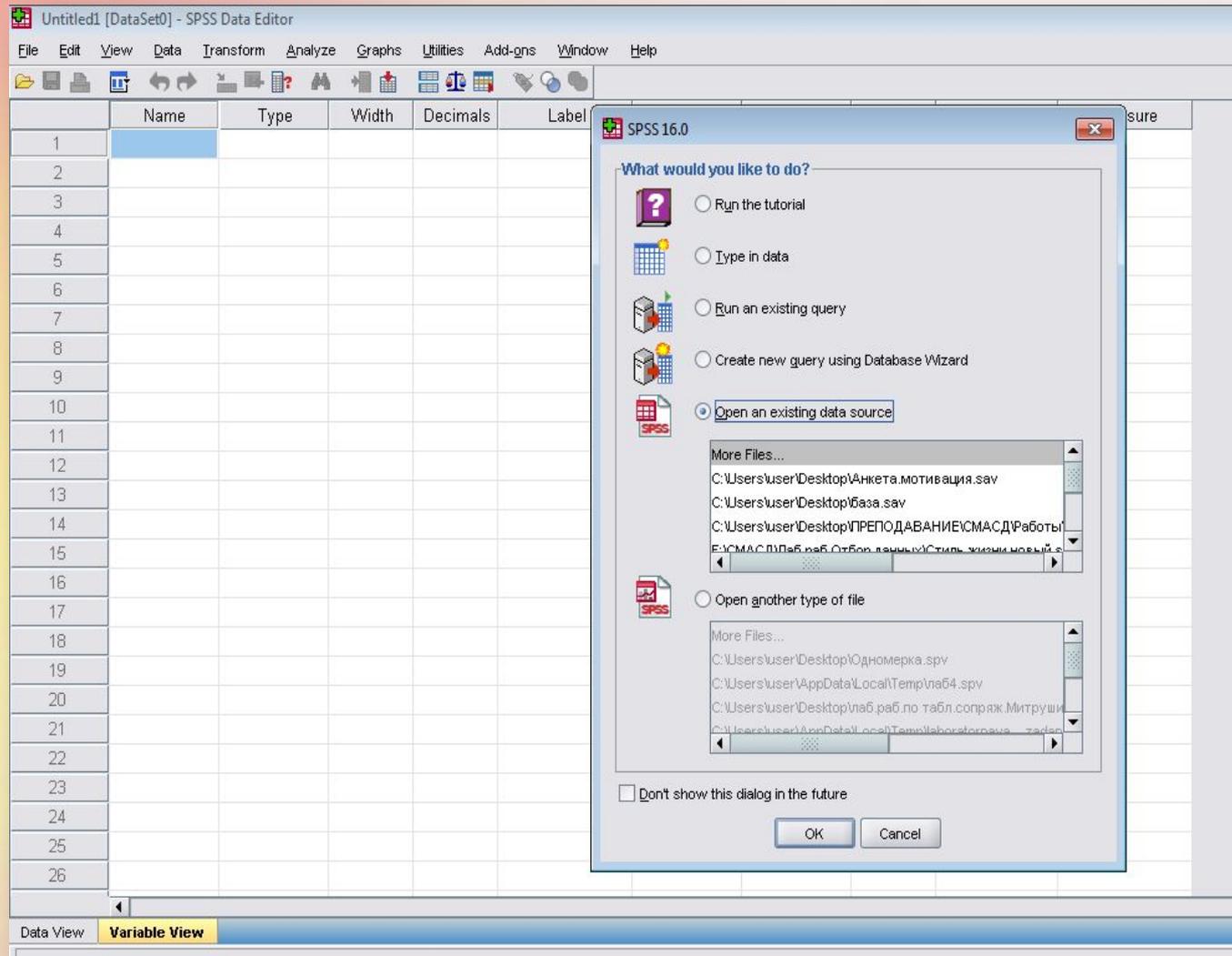
В SPSS реализованы следующие методы статистической обработки информации:

- суммарные статистики по отдельным переменным;
- частоты, суммарные статистики и графики для произвольного числа переменных;
- построение N-мерных таблиц сопряженности и получение мер связи;
- средние, стандартные отклонения и суммы по группам;
- дисперсионный анализ и множественные сравнения;
- корреляционный анализ;
- дискриминантный анализ;
- однофакторный дисперсионный анализ;
- факторный анализ;
- кластерный анализ;
- многомерный дисперсионный анализ;
- непараметрические тесты;
- множественная регрессия;
- методы оптимального шкалирования и т.д.

Запуск SPSS. Редактор данных

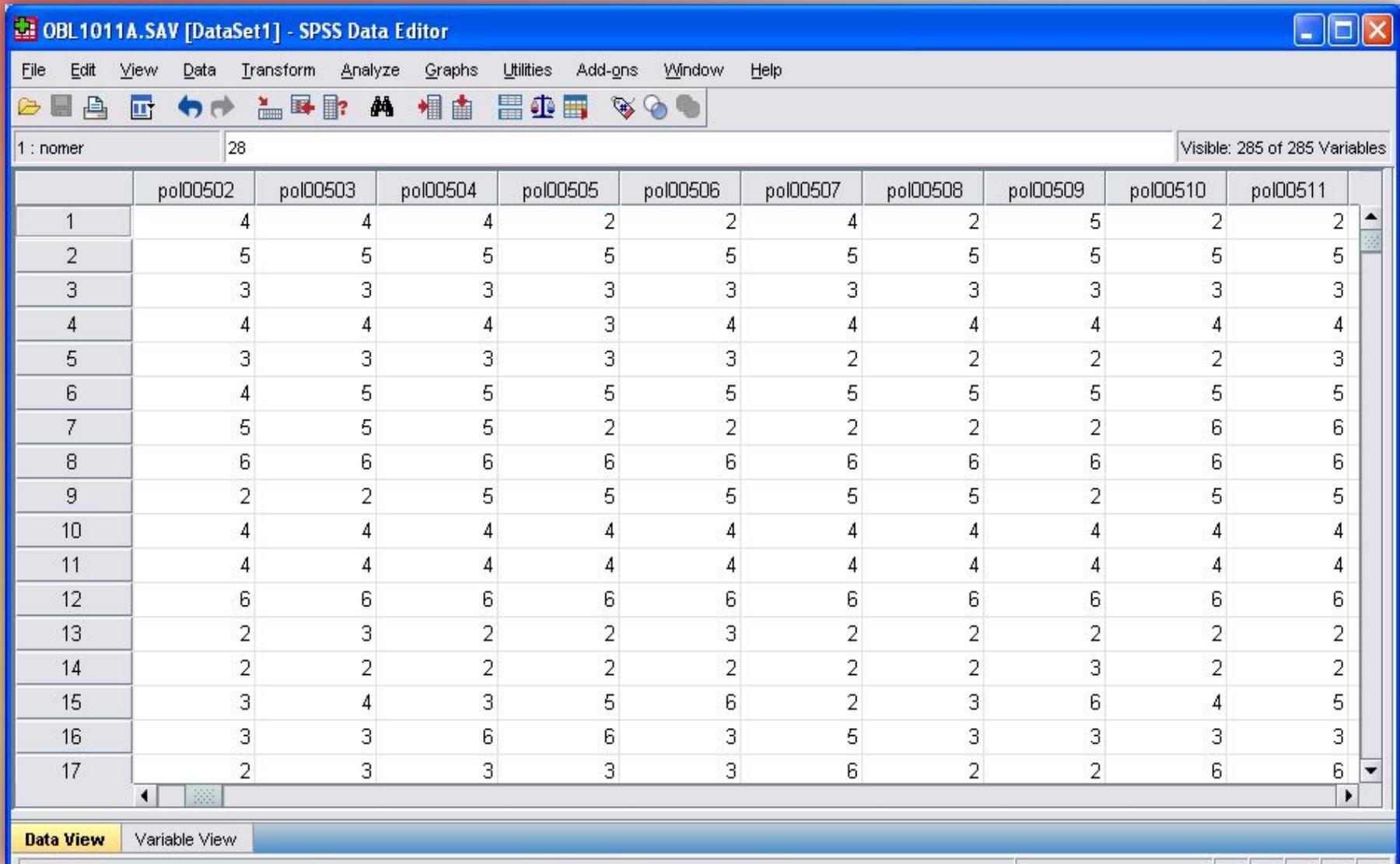
Редактор данных – это одно из многих окон SPSS. Здесь можно вводить новые данные или загружать существующие из файлов данных с помощью команд меню **File (Файл) > Open... (Открыть...)**.

Так как при запуске SPSS ни один файл данных еще не загружен, в заголовке редактора данных стоит "**Untitled**" (Без имени). Над изображением таблицы в редакторе данных имеются строка меню и панель символов.



Запуск SPSS. Редактор данных

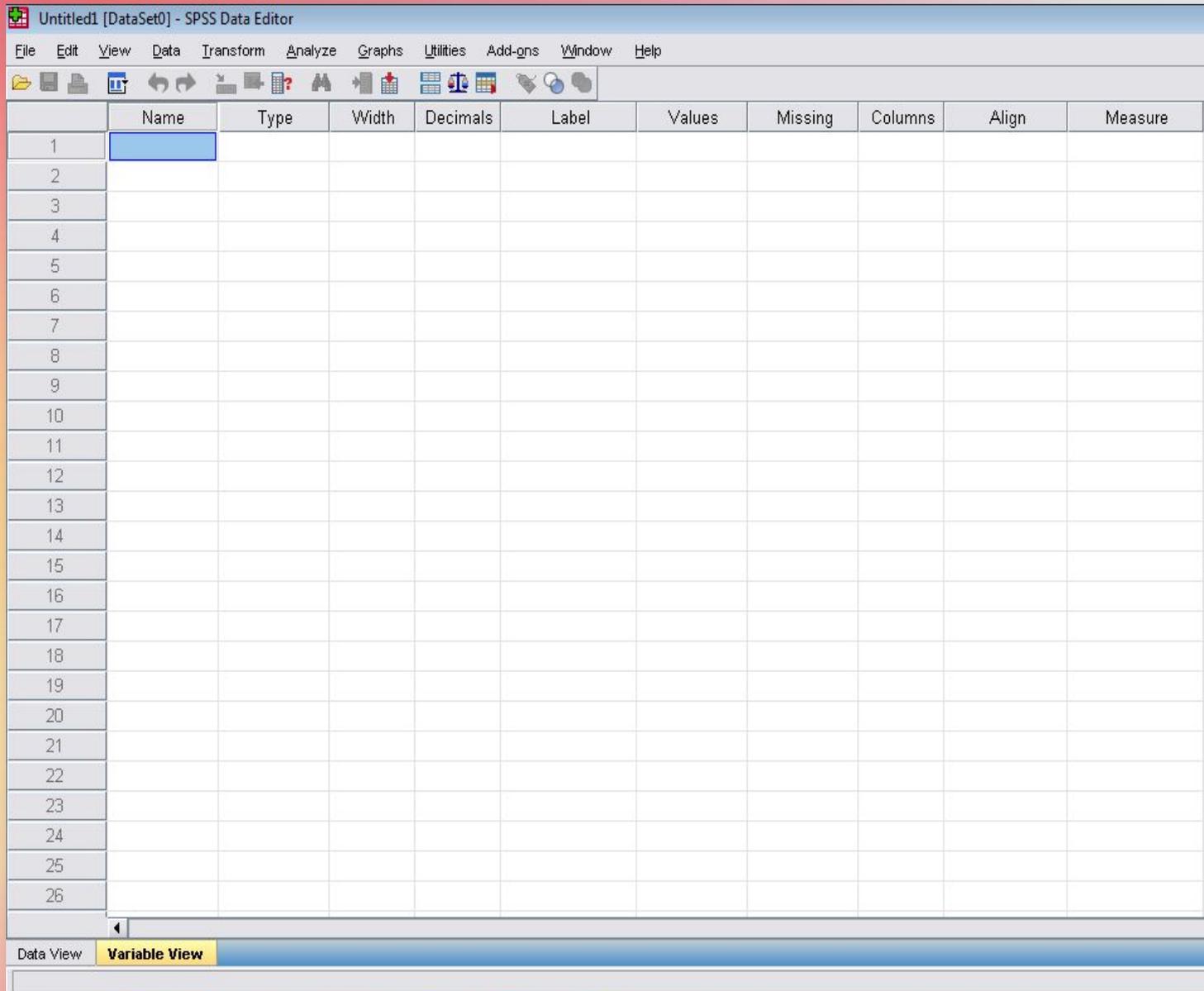
Вид листа ввода переменных



The screenshot shows the SPSS Data Editor window for a file named 'OBL1011A.SAV [DataSet1]'. The window title bar includes standard Windows window controls. The menu bar contains 'File', 'Edit', 'View', 'Data', 'Transform', 'Analyze', 'Graphs', 'Utilities', 'Add-ons', 'Window', and 'Help'. Below the menu bar is a toolbar with various icons for file operations and data manipulation. The main area displays a data table with 17 rows and 11 columns. The first column is labeled '1: nomer' and contains row numbers from 1 to 17. The other columns are labeled 'pol00502' through 'pol00511'. The data values are integers ranging from 2 to 6. At the bottom of the window, there are tabs for 'Data View' (which is selected) and 'Variable View'.

	pol00502	pol00503	pol00504	pol00505	pol00506	pol00507	pol00508	pol00509	pol00510	pol00511
1	4	4	4	2	2	4	2	5	2	2
2	5	5	5	5	5	5	5	5	5	5
3	3	3	3	3	3	3	3	3	3	3
4	4	4	4	3	4	4	4	4	4	4
5	3	3	3	3	3	2	2	2	2	3
6	4	5	5	5	5	5	5	5	5	5
7	5	5	5	2	2	2	2	2	6	6
8	6	6	6	6	6	6	6	6	6	6
9	2	2	5	5	5	5	5	2	5	5
10	4	4	4	4	4	4	4	4	4	4
11	4	4	4	4	4	4	4	4	4	4
12	6	6	6	6	6	6	6	6	6	6
13	2	3	2	2	3	2	2	2	2	2
14	2	2	2	2	2	2	2	3	2	2
15	3	4	3	5	6	2	3	6	4	5
16	3	3	6	6	3	5	3	3	3	3
17	2	3	3	3	3	6	2	2	6	6

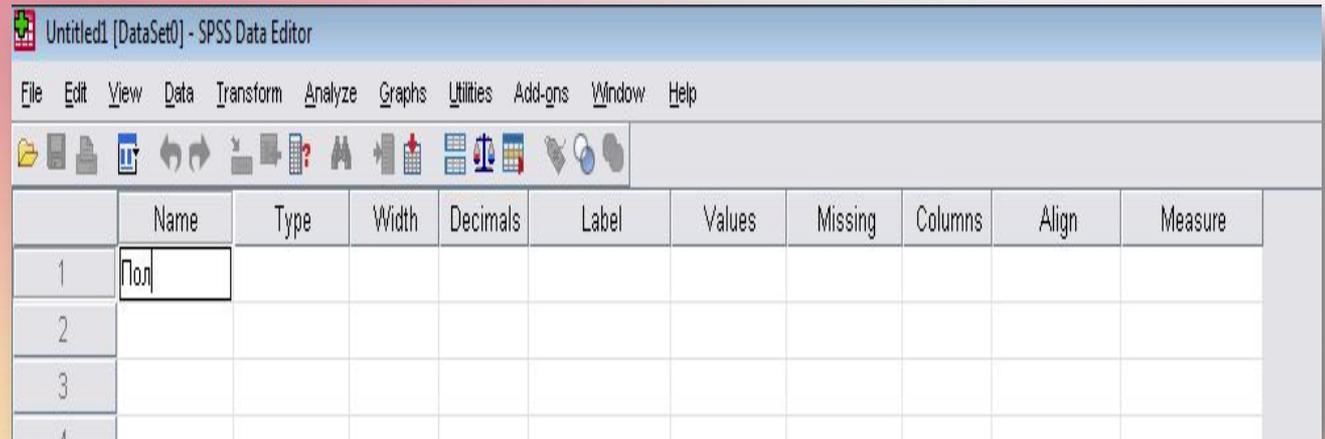
Определение переменных



В редакторе данных дважды щелкните на ячейке с надписью **var** или щелкните на ярлычке **Variable view** (Просмотр переменных) на нижнем краю таблицы.

Имя переменной

Чтобы задать имя переменной, необходимо:
Ввести в текстовом поле **Name** (Имя) выбранное имя переменной.



- При выборе имени переменной следует соблюдать определенные правила:*
1. Имена переменных могут содержать буквы латинского алфавита и цифры. Кроме того, допускаются специальные символы _ (подчеркивание), . (точка), а также символы @ и #. Не разрешаются, например, пробелы, знаки других алфавитов и специальные символы, такие как !, ?, " и *.
 2. Имя переменной должно начинаться с буквы.
 3. Последний символ имени не может быть точкой или знаком подчеркивания (_).
 3. Длина имени переменной **не должна превышать восьми символов.**
 4. Имена переменных нечувствительны к регистру, то есть прописные и строчные буквы не различаются.

Имя переменной



Примеры допустимых имен переменных:

пол

зарплата

возраст

экпол_2

Примеры недопустимых имен переменных:

1na1 Имя начинается не с буквы

Assignment Имя длиннее 8 символов

State 94 Имя содержит пробел

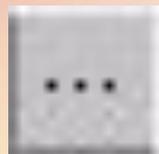
None! Символ "!" не разрешается



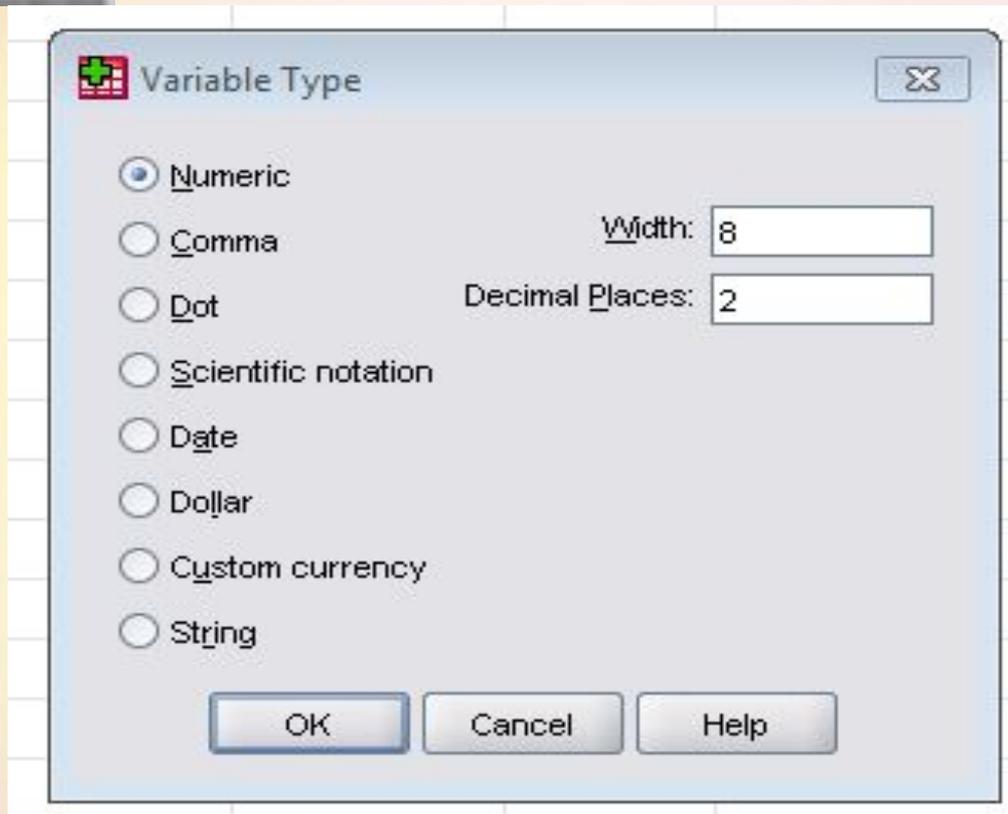
Тип переменных

Вновь созданные в SPSS переменные по умолчанию являются численными с максимальной длиной восемь знаков, причем дробная часть состоит из двух знаков (формат F8.2).

- Если требуется изменить тип переменной, щелкните в ячейке на кнопке с тремя точками:

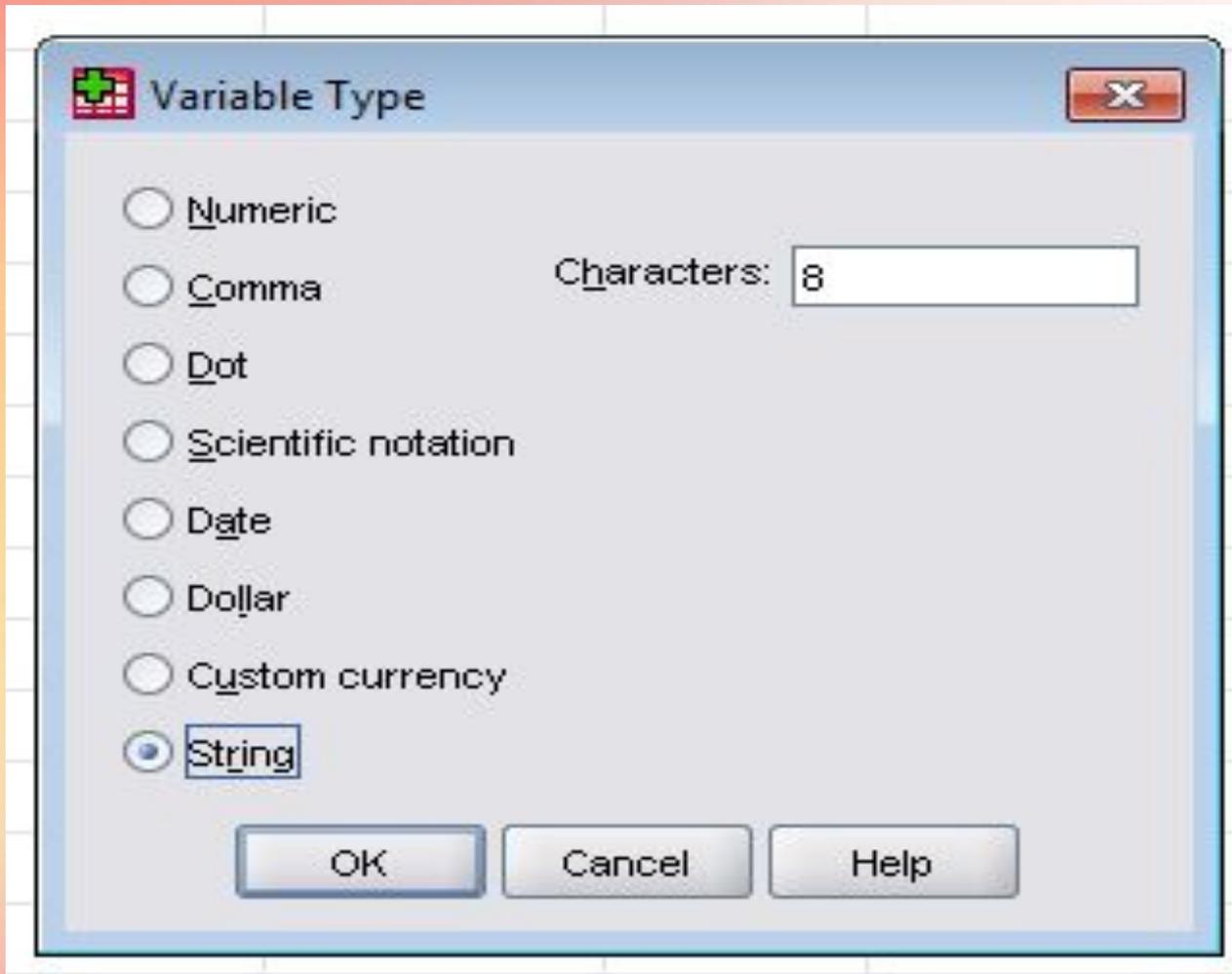


Откроется
диалоговое окно
**Define Variable
Type** (Определение
типа переменной).



Тип переменных

Строчные (String) переменные используются достаточно редко, например, для введения ответов на открытые вопросы или фамилий респондентов,



С такими переменными нельзя выполнять никаких вычислительных операций, но можно проводить, например, подсчеты повторяемости.

Типы переменных

Наименование	Характеристика
Numeric (Численный)	К допустимым значениям относятся цифры, перед которыми стоит знак плюс или минус и десятичный разделитель. Знак плюс перед числом, в отличие от минуса, не отображается. В текстовом поле Length (Длина) задается максимальное количество знаков, включая позицию для десятичного разделителя. В текстовом поле Decimals (Десятичные разряды) вводится количество отображаемых знаков дробной части.
Comma (Запятая)	К допустимым значениям относятся цифры, перед которыми стоит знак плюс или минус, точка, как десятичный разделитель и одна или несколько запятых в качестве разделителей групп разрядов. Если запятые опускаются при вводе, они вставляются автоматически. Длина такой переменной равна максимальному количеству знаков, включая десятичный разделитель и запятые между группами разрядов.

Типы переменных

Наименование	Характеристика
Dot (Точка)	К допустимым значениям относятся цифры, перед которыми стоит знак плюс или минус, запятая, как десятичный разделитель и одна или несколько точек в качестве разделителей групп разрядов. Если точки опускаются при вводе, они вставляются автоматически.
Scientific notation (Экспоненциальное представление)	При вводе данных разрешаются все допустимые численные значения, включая экспоненциальное представление, о котором свидетельствует содержащаяся в числе буква E или D, а также знак плюс или минус.
Date (Дата)	Допустимые значения – дата и/или время.

Типы переменных

Наименование	Характеристика
Dollar (Доллар)	К допустимым значениям относятся: знак доллара, точка, как десятичный разделитель и запятые, как разделители групп разрядов. Если знак доллара или запятые опускаются при вводе, они вставляются автоматически.
Special currency (Специальная валюта)	Пользователь может задавать собственные форматы валюты. В поле Length в этом случае задается максимальное количество знаков, включая все знаки, заданные пользователем. Обозначение валюты при вводе не указывается; оно вставляется автоматически.
String (Строка)	Строка символов. К допустимым значениям относятся: буквы, цифры и специальные символы. Различаются короткие и длинные строковые переменные. Короткие строковые переменные могут содержать не более восьми знаков. В большинстве процедур SPSS применение длинных строковых переменных ограничивается или вообще не допускается.

Метка переменной и метки значений

Метка переменной (Label) – это название, позволяющая описать переменную более подробно.

Метка переменной может содержать до 256 символов.

В метках переменных различаются прописные и строчные буквы. Они отображаются в том виде, в каком были введены.



Метки значений (Values) – это название, позволяющее более подробно описать возможные значения переменной.

Так, например, в случае переменной Пол можно задать метку " мужской " для значения "1" и метку «женский" для значения "2". Подтвердить настройку по умолчанию **None** (Нет) нужно клавишей **Tab**. Впрочем, ввод данных также можно подтвердить клавишей **Enter**.

Пропущенные значения (*Missing values*)

В SPSS допускаются два вида пропущенных значений:

1. Пропущенные значения, определяемые системой (**System-defined missing values**): Если в матрице данных есть незаполненные численные ячейки, система SPSS самостоятельно идентифицирует их как пропущенные значения. Этот факт отображается в матрице данных с помощью запятой (,).

2. Пропущенные значения, задаваемые пользователем (**User-defined missing values**): Если в определенных случаях у переменных отсутствуют значения, например, если на вопрос не был дан ответ, ответ неизвестен, или существуют другие причины, пользователь может с помощью кнопки **Missing** объявить эти значения как пропущенные. Пропущенные значения можно исключить из последующих вычислений.

Нужно подтвердить настройку по умолчанию **None** (Нет) клавишей **Enter**.



Столбцы (*Columns*) и Выравнивание (*Alignment*)

Поле *Columns* определяет ширину, которую будет иметь в таблице данный столбец при отображении значений. Ширину столбца также можно изменить непосредственно в окне редактора данных. Для этого поместите указатель мыши на разделитель между двумя заголовками столбцов с именами переменных. Вид указателя изменится. Появившаяся двойная стрелка указывает, что соответствующий столбец можно расширить или сузить путем перетаскивания. **Нужно** подтвердить настройку по умолчанию "8" клавишей **Enter**.



Выравнивание (*Alignment*):

Здесь можно задать вид выравнивания значений, т.е. определить, как они будут отображаться в таблице. Возможные виды выравнивания – "**Right**" (по правому краю), "**Left**" (по левому краю) и "**Center**" (по центру). Чтобы задать вид выравнивания, щелкните на кнопке



Нужно подтвердить настройку по умолчанию **Right** клавишей **Enter**.

Width (ширина) и *Decimals* (десятичные значения)



Width (ширина) – максимальное количество знаков, которое имеет переменная, включая дробную часть.



Decimals (десятичные значения) – количество знаков после запятой.

Шкала измерения (*Measure*)

Здесь можно задать шкалу переменной, которая может быть номинальной (шкала наименований), порядковой или метрической.

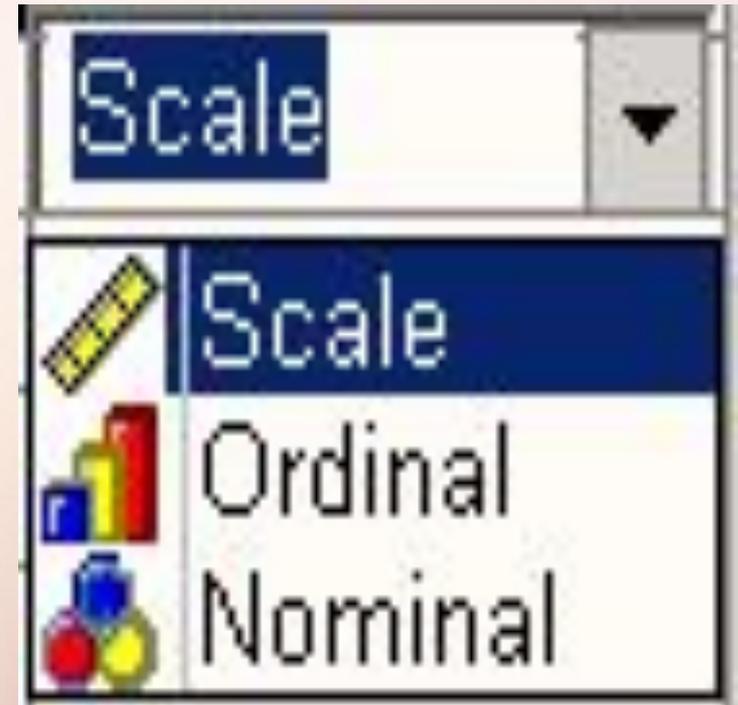
По умолчанию принимается метрическая шкала измерения.

Правда, это различие имеет значение только при создании интерактивных графиков, где номинальная и порядковая шкала измерений объединяются в "категориальный" тип.

Если загружаются файлы, созданные в предыдущих версиях SPSS, или шкала измерений не определяется явно, SPSS вначале автоматически предполагает метрическую шкалу. Однако, если соответствующая переменная имеет метки значений или принимает менее 24 различных значений, то задается порядковая шкала.

Нужно: подтвердить настройку по умолчанию **Nominal** (шкала наименований) клавишей **Tab**.

Затем снова поместить курсор в поле **Name**, чтобы начать объявление следующей переменной.



Основные команды меню SPSS

FILE: обеспечивает доступ к файлам данных, к выходным файлам и программам преобразования данных.

EDIT: редактирование командных файлов, выходных файлов и файлов данных статистических наблюдений и др.

DATA: обеспечивает операции над данными: сортировку, слияние файлов данных, агрегирование, организацию подвыборки из данных.



TRANSFORM: преобразование данных.

ANALIS: реализация методов анализа данных

GRAPHS: графическое представление данных.

UTILITIES: обслуживающие программы.

WINDOOW: переключение окон.

HELP: справочная информация.

Пример создания переменной

The screenshot shows the SPSS Data Editor interface. The main window displays a variable definition table for a variable named 'Пол' (Sex). The table has columns for Name, Type, Width, Decimals, Label, Values, Missing, Columns, Align, and Measure. The 'Values' column for 'Пол' is currently set to ', мужской}...'.

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	Пол	Numeric	8	0	Ваш пол	, мужской}...	None	8	Right	Scale
2										
3										
4										
5										
6										
7										
8										
9										
10										
11										
12										
13										
14										
15										
16										
17										
18										
19										
20										
21										
22										
23										
24										
25										
26										

The 'Value Labels' dialog box is open, showing the following configuration:

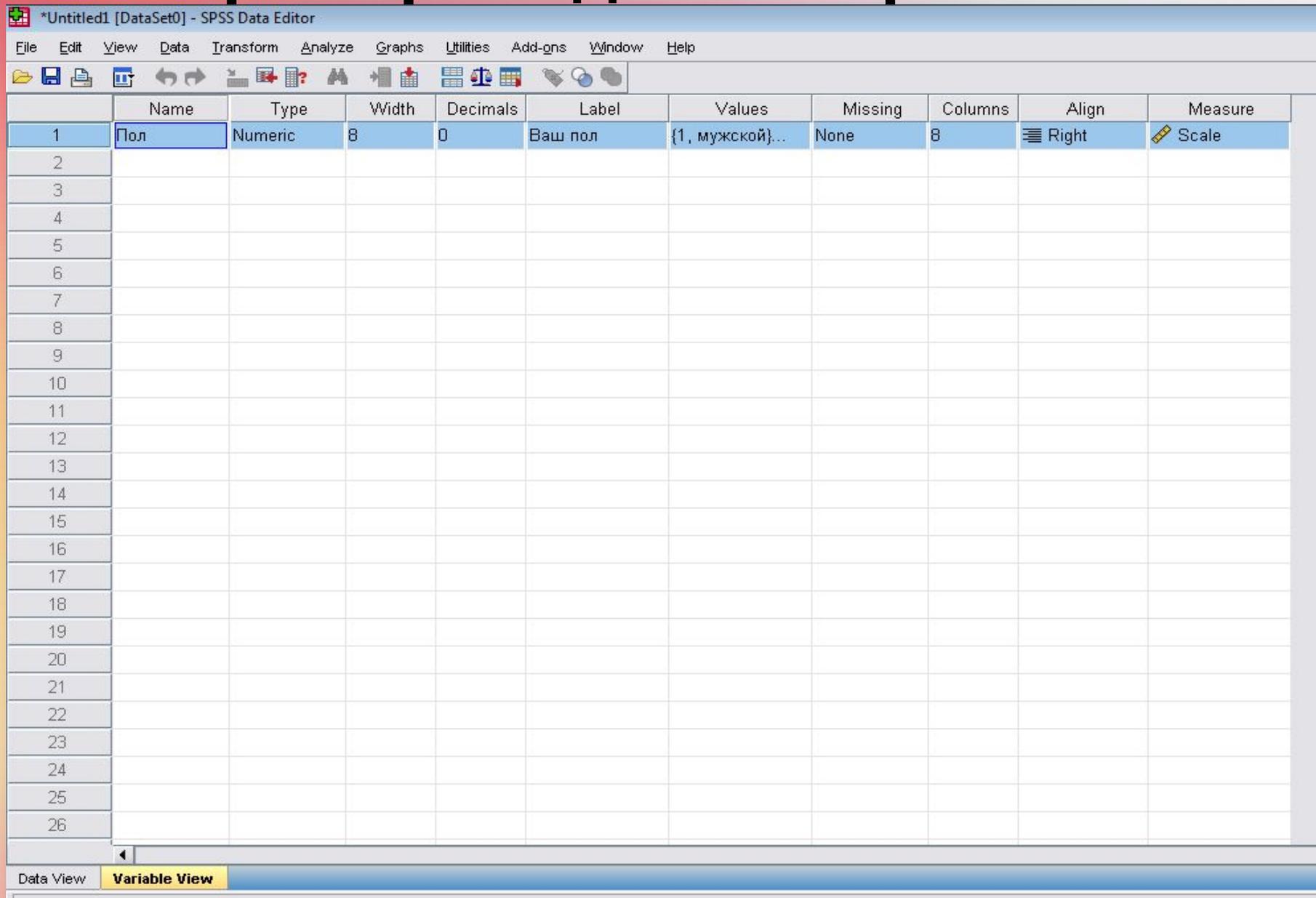
- Value:
- Label:
- Buttons: Add, Change, Remove, Spelling...
- List of labels: 1 = "мужской", 2 = "женский"
- Buttons: OK, Cancel, Help

The bottom of the window shows the 'Variable View' tab selected.

Пример создания переменной

*Untitled1 [DataSet0] - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help



The screenshot shows the SPSS Data Editor interface in Variable View. The main window displays a table with columns for variable properties and a list of variables on the left. The first variable, 'Пол', is highlighted in blue. The table columns are: Name, Type, Width, Decimals, Label, Values, Missing, Columns, Align, and Measure. The 'Пол' variable is defined as Numeric, 8 digits wide, 0 decimal places, with the label 'Ваш пол' and values '{1, мужской}...'. It is set to 'None' for missing values, 8 columns wide, right-aligned, and measured as a 'Scale'.

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	Пол	Numeric	8	0	Ваш пол	{1, мужской}...	None	8	Right	Scale
2										
3										
4										
5										
6										
7										
8										
9										
10										
11										
12										
13										
14										
15										
16										
17										
18										
19										
20										
21										
22										
23										
24										
25										
26										

Data View Variable View

Пример создания переменной

The screenshot shows the SPSS Data Editor interface. The title bar reads '*Untitled1 [DataSet0] - SPSS Data Editor'. The menu bar includes File, Edit, View, Data, Transform, Analyze, Graphs, Utilities, Add-ons, Window, and Help. The toolbar contains various icons for file operations and data manipulation. The main window displays a data grid with 25 rows and 17 columns. The first column is labeled 'Пол' and contains the values 1, 2, 1, 2, 1, 1, 1, 2, 2, 2, followed by empty cells for rows 11 through 25. The other 16 columns are labeled 'var'. The status bar at the bottom indicates 'SPSS Processor is ready'.

	Пол	var															
1	1																
2	2																
3	1																
4	2																
5	1																
6	1																
7	1																
8	2																
9	2																
10	2																
11																	
12																	
13																	
14																	
15																	
16																	
17																	
18																	
19																	
20																	
21																	
22																	
23																	
24																	
25																	

**Спасибо за
внимание!**