

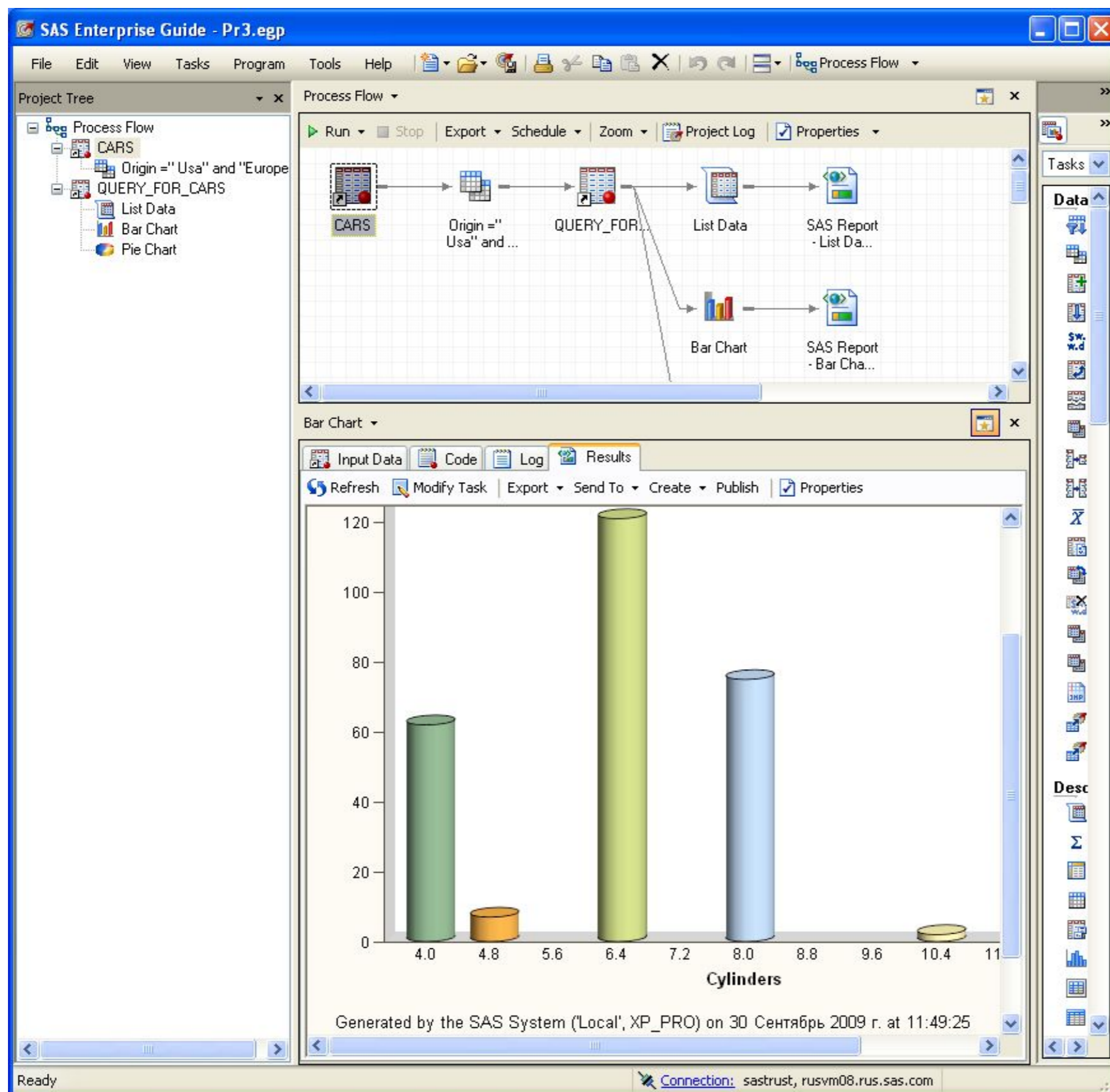
День 1

Раздел 1. Введение в SAS Enterprise Guide

1.1. Обзор интерфейса

1.2. Рабочие области проектов, обновление результатов

Знакомство с SAS Enterprise Guide



Знакомство SAS Enterprise Guide

- SAS Enterprise Guide представляет удобный графический интерфейс для работы с данными и создания отчетов

The image shows two overlapping dialog boxes in SAS Enterprise Guide. The background dialog is titled 'Bar Chart for Local:SASUSER.QUERY_FOR_CARS' and is in the 'Data' tab. It shows the data source as 'Local:SASUSER.QUERY_FOR_CARS' and a list of columns to assign, including Make, Model, Type, Origin, DriveTrain, MSRP, Invoice, EngineSize, Cylinders, Horsepower, MPG_City, MPG_Highway, Weight, and Wheelbase. The foreground dialog is titled 'Bar Chart1 for Local:SASHELP.CARS' and is in the 'Specify appearance' step (3 of 4). It shows options for 3D chart, colors, labels, and ticks. A sample 3D bar chart is displayed on the right, showing the frequency of cylinders (1, 2, 3).

Bar Chart1 for Local:SASHELP.CARS - Specify appearance (3 of 4)

Options:

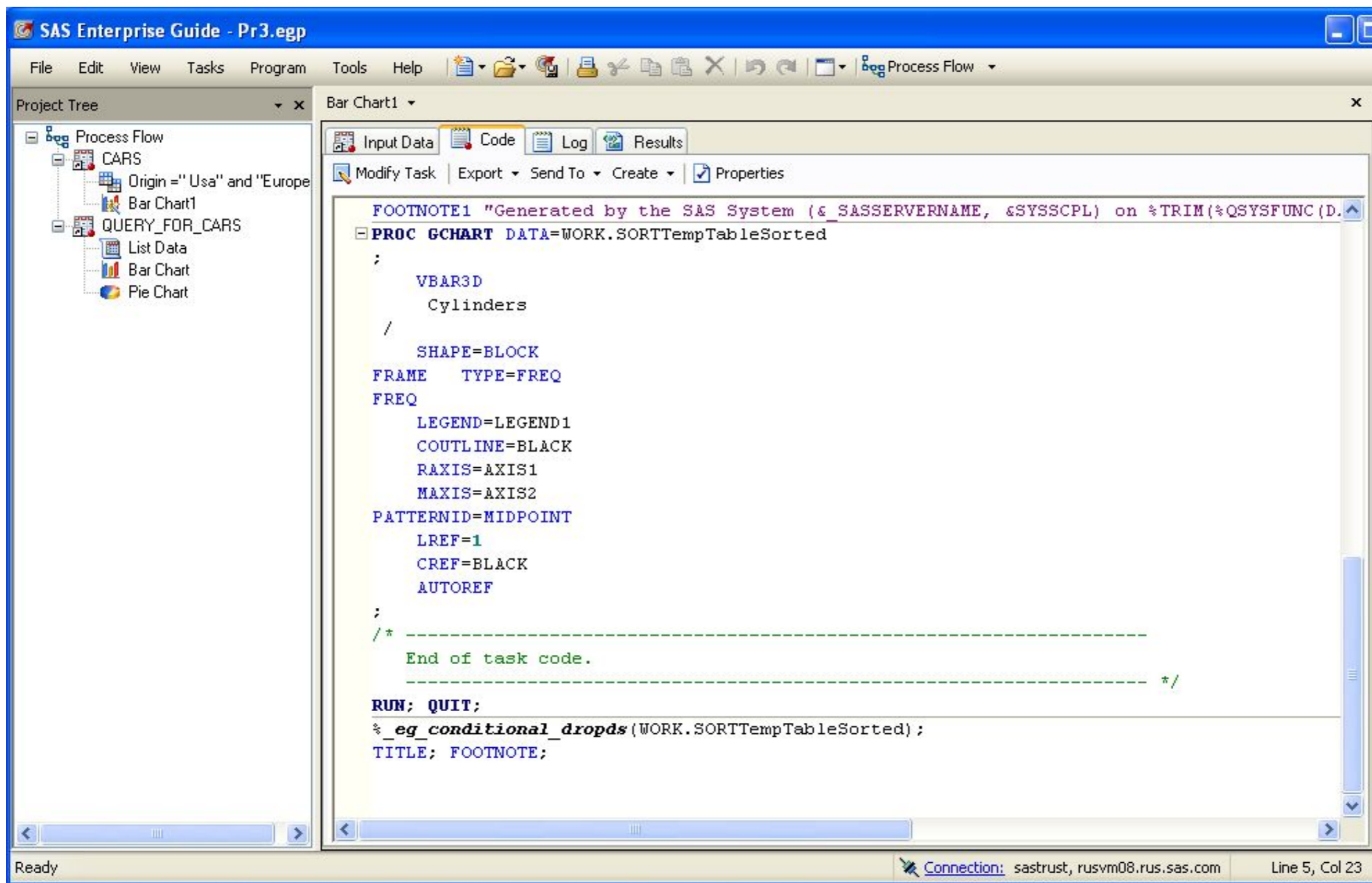
- 3D chart
- Colors: Color bars by: Bar category
- Labels: Legend: South; Data labels: Frequency
- Lines and ticks: Use reference lines; Tick marks

Sample chart: (Frequency)

Cylinders	Frequency
1	1
2	1
3	1

SAS Enterprise Guide Интерфейс

- SAS Enterprise Guide также включает в себя программный интерфейс, позволяющий создавать, редактировать, и запускать SAS код.



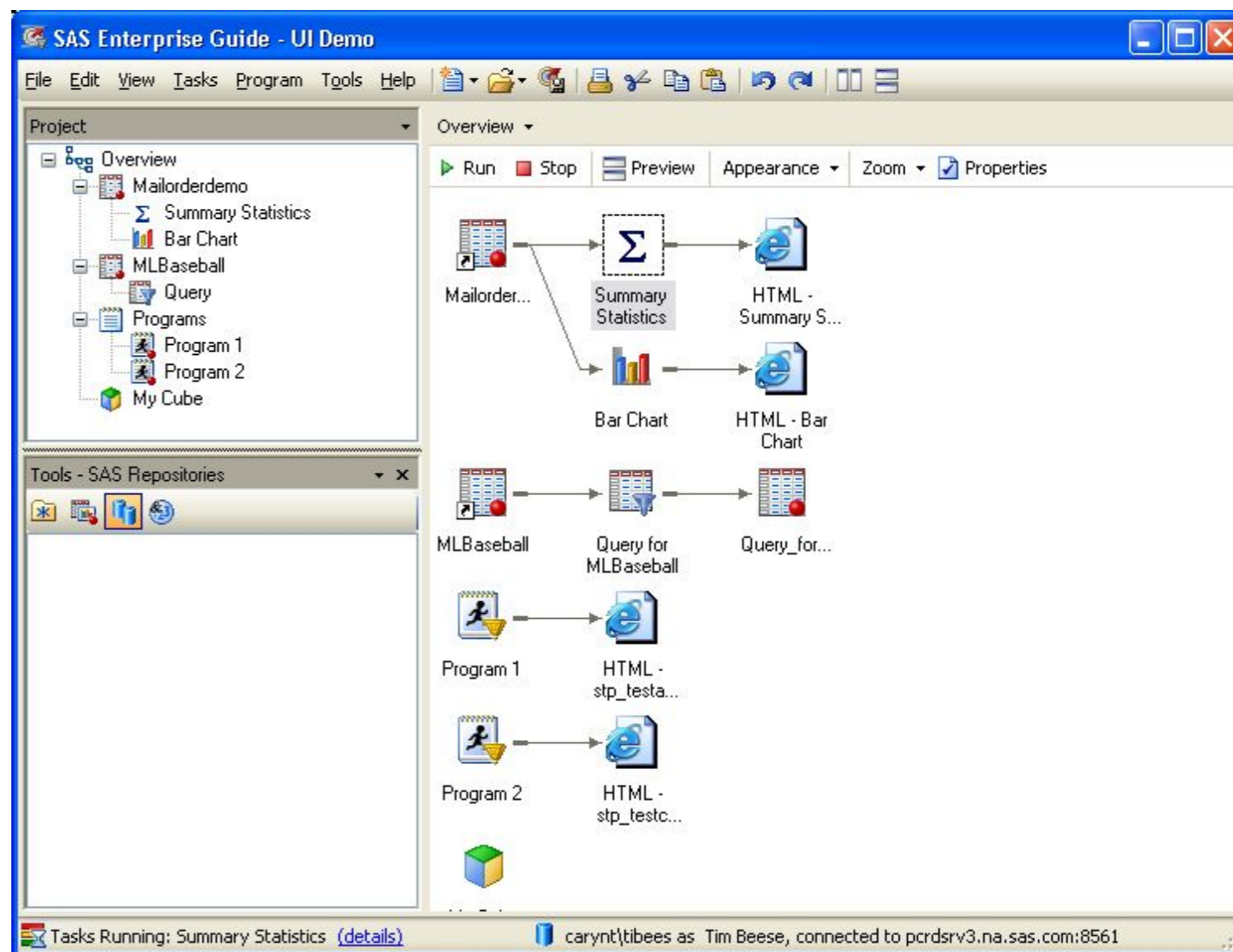
The screenshot displays the SAS Enterprise Guide interface. The main window is titled "SAS Enterprise Guide - Pr3.egg". The menu bar includes File, Edit, View, Tasks, Program, Tools, and Help. The Project Tree on the left shows a process flow with tasks: CARS, Origin = " Usa" and "Europe", Bar Chart1, QUERY_FOR_CARS, List Data, Bar Chart, and Pie Chart. The main editor window is titled "Bar Chart1" and contains the following SAS code:

```
FOOTNOTE1 "Generated by the SAS System (&_SASSERVERNAME, &SYSSCPL) on %TRIM(%QSYFUNC(D.
PROC GCHART DATA=WORK.SORTTempTableSorted
;
  VBAR3D
    Cylinders
  /
  SHAPE=BLOCK
  FRAME TYPE=FREQ
  FREQ
  LEGEND=LEGEND1
  COUTLINE=BLACK
  RAXIS=AXIS1
  MAXIS=AXIS2
  PATTERNID=MIDPOINT
  LREF=1
  CREF=BLACK
  AUTOREF
;
/* -----
End of task code.
----- */
RUN; QUIT;
%_eg_conditional_drops(WORK.SORTTempTableSorted);
TITLE; FOOTNOTE;
```

The status bar at the bottom shows "Ready", "Connection: sastrust, rusvm08.rus.sas.com", and "Line 5, Col 23". The "byte consulting company" logo is visible in the bottom right corner.

SAS Enterprise Guide: Рабочая область проекта

- Проект это набор
 - Источников данных
 - SAS программ и логов
 - Аналитических задач и запросов
 - Результатов
 - Записок и вспомогательной документации



Вы можете также управлять содержимым, изменять последовательность исполнения узлов и порядок обновления результатов в проекте.

Создание нового проекта в SAS Enterprise Guide

Для того чтобы начать работу с SAS Enterprise Guide, нужно

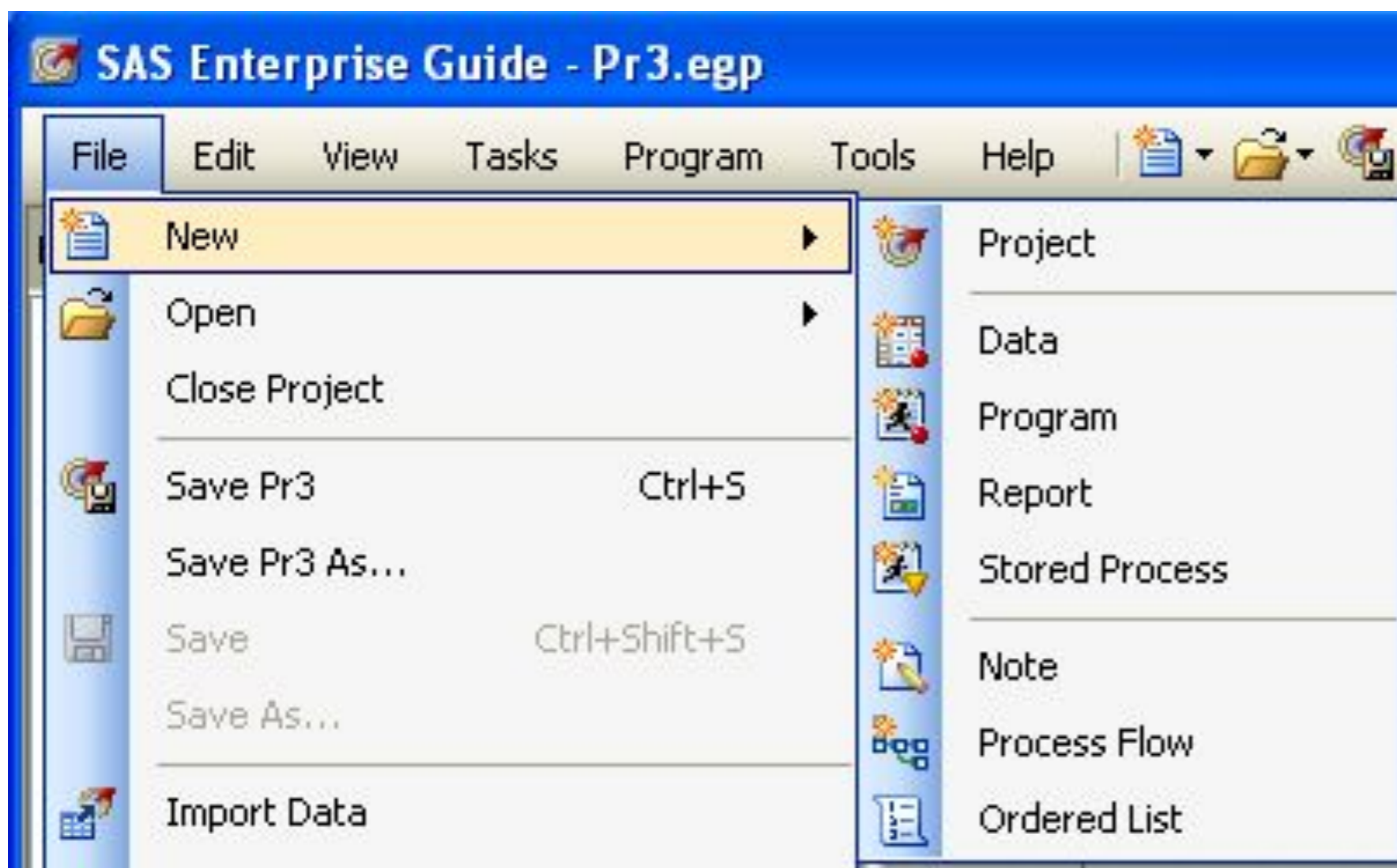
1. Создать новый проект
2. Добавить данные в проект
3. Запустить аналитическую задачу на данных.

Также вы можете

4. Кастомизировать результаты
5. Автоматизировать процесс выполнения проекта.
6. Вложить некоторую логику в последовательность узлов

Организация рабочих областей в проекте

- Для больших проектов, бывает полезно создавать более одной рабочей области в проекте для того чтобы группировать схожие запросы и задачи. Нажать кнопку **File** -> **New** -> **Process Flow** для того чтобы добавить новую рабочую область в проект.



Создание новых рабочих областей проекта

- В рамках проекта SAS EG можно создавать несколько рабочих областей, что позволяет удобней управлять содержимым проекта и обновлять задачи и запросы в проекте, кроме того можно изменять вид отображения выходных окон.

The screenshot displays the SAS Enterprise Guide (EG) interface. The title bar reads "SAS Enterprise Guide - EG_demo_6.egp". The menu bar includes "File", "Edit", "View", "Tasks", "Program", "Tools", and "Help". A toolbar contains icons for file operations and a "Process Flow" dropdown menu. The "Process Flow" menu is open, showing options: "Single", "Stacked", and "Side By Side".

The main workspace shows a process flow diagram on a grid. The flow starts with a task named "AUTOMPG" (represented by a data table icon). From "AUTOMPG", two arrows branch out: one to a "Bar Chart" task and another to a "Pie Chart" task. The "Bar Chart" task is connected to a "SAS Report - Bar Cha..." task, and the "Pie Chart" task is connected to a "SAS Report - Pie Cha..." task. Below the main workspace, there is a "Task Status" window with a table structure:

Task	Status	Queue
------	--------	-------

Раздел 2: Работа с данными в проекте

2.1. Доступ к локальным данным

2.2. Доступ к удаленным источникам данных (Внешним СУБД)

2.3. Формат исходных данных

2.4. Импорт из текстовых файлов

2.5. Экспорт в файл

Общие форматы данных

SAS Enterprise Guide позволяет пользователю работать с любыми форматами данных:

- 1) Таблицы внешних СУБД, ODBC, Oracle, Sybase, MS Access
- 2) SAS таблицы, Excel 2003, 2007, 2010 таблицы, HTML таблицы
- 3) Текстовые файлы с разделителями и фиксированной длины.
- 4) SPSS , Stata файлы, файлы других вендоров .

Для того чтобы, настроить внешние данные, необходимо либо:

- Определить библиотеку с помощью кода в проекте, либо в startup code (запускном bat файле).
- Определить библиотеку и доступ к внешним данным, на сервере метаданных SAS, с помощью SAS Management Console.

Открытие удаленных данных

- Иконки серверов в SAS EG позволяют получить доступ к удаленным данным расположенных на серверах SAS, и прочим объявленным на сервере метаданным внешним СУБД

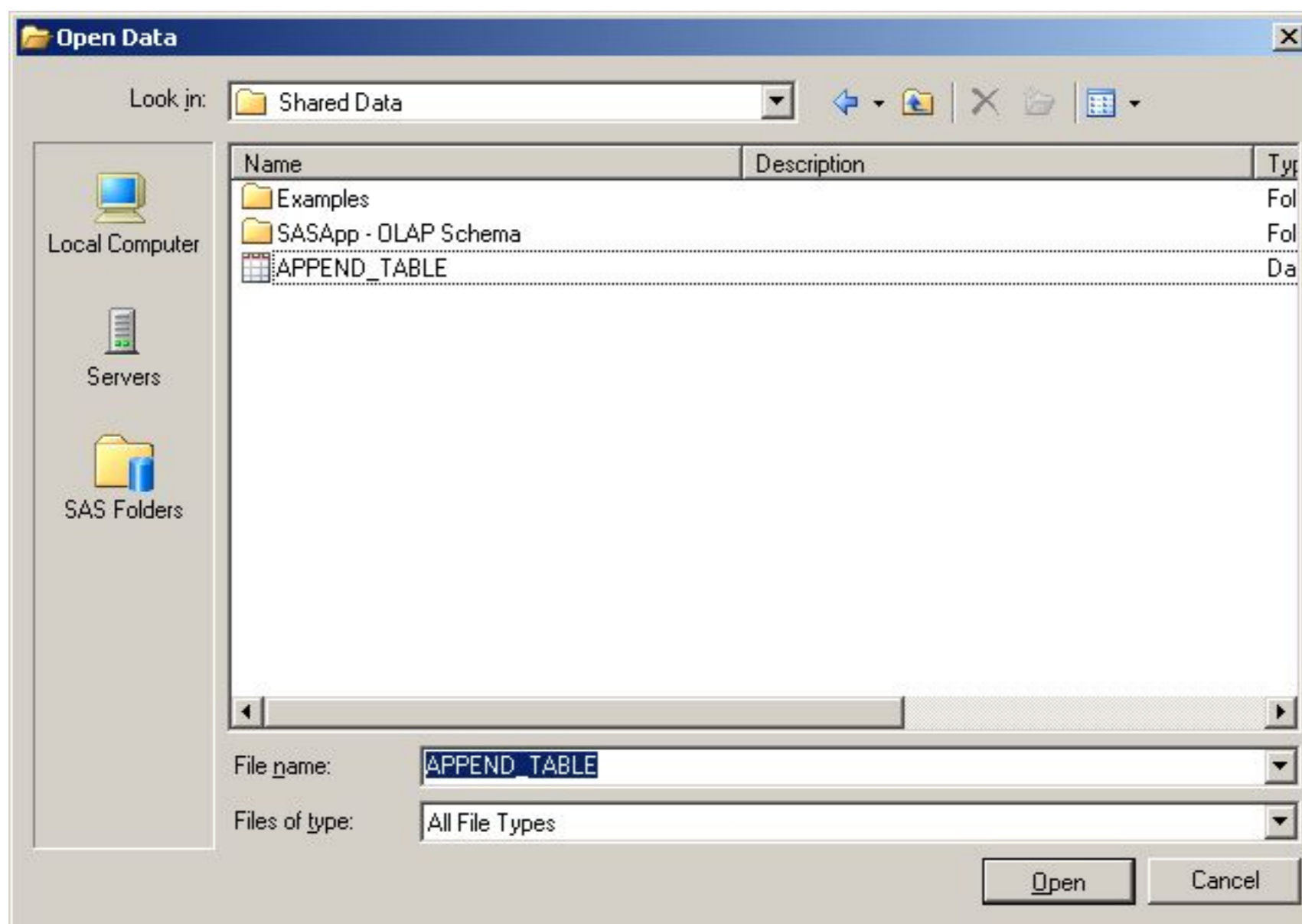


Иконка Libraries позволяет получить доступ к SAS папкам или Внешним СУБД.

Папка с файлами дает доступ к директориям на сервере

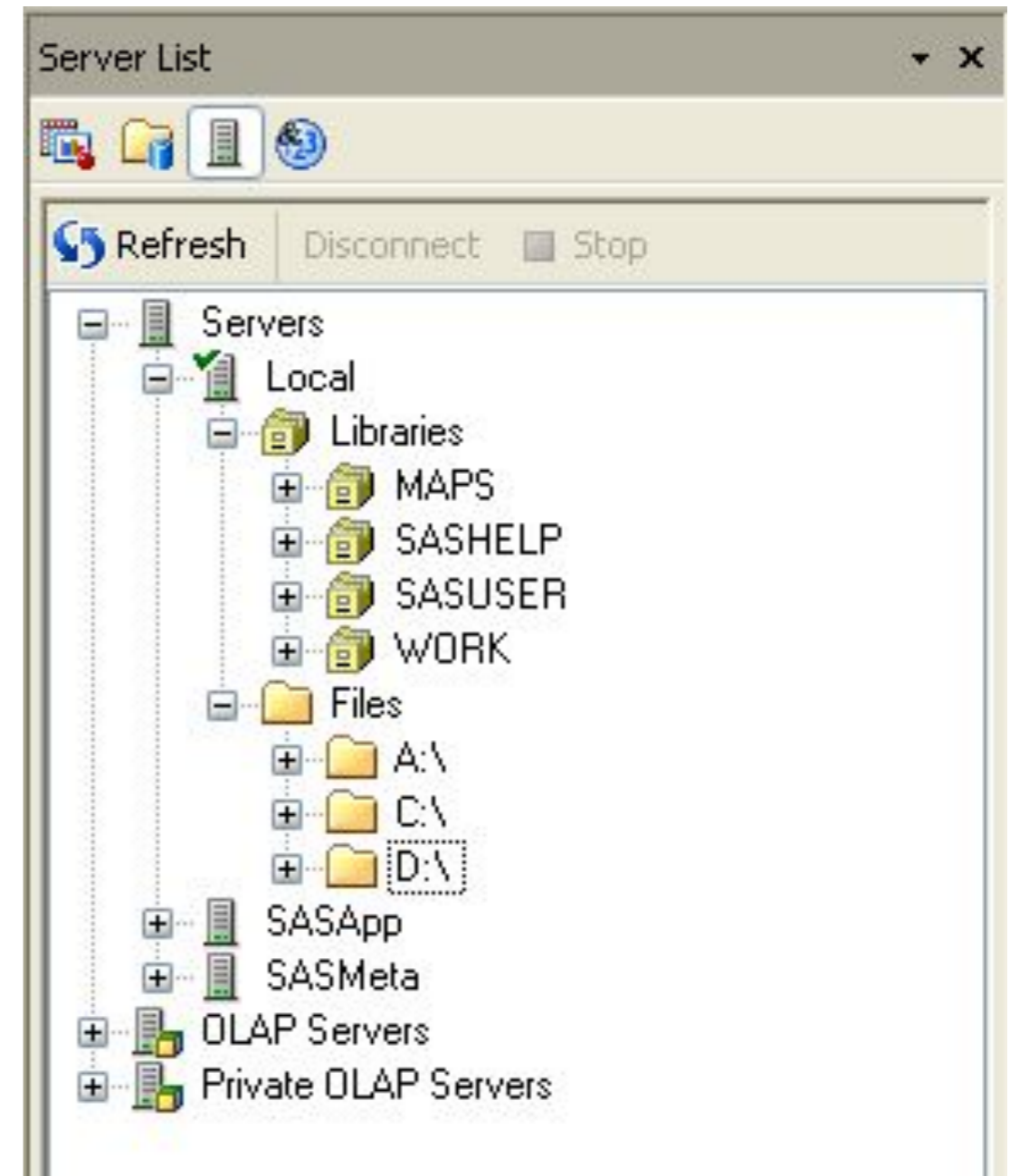
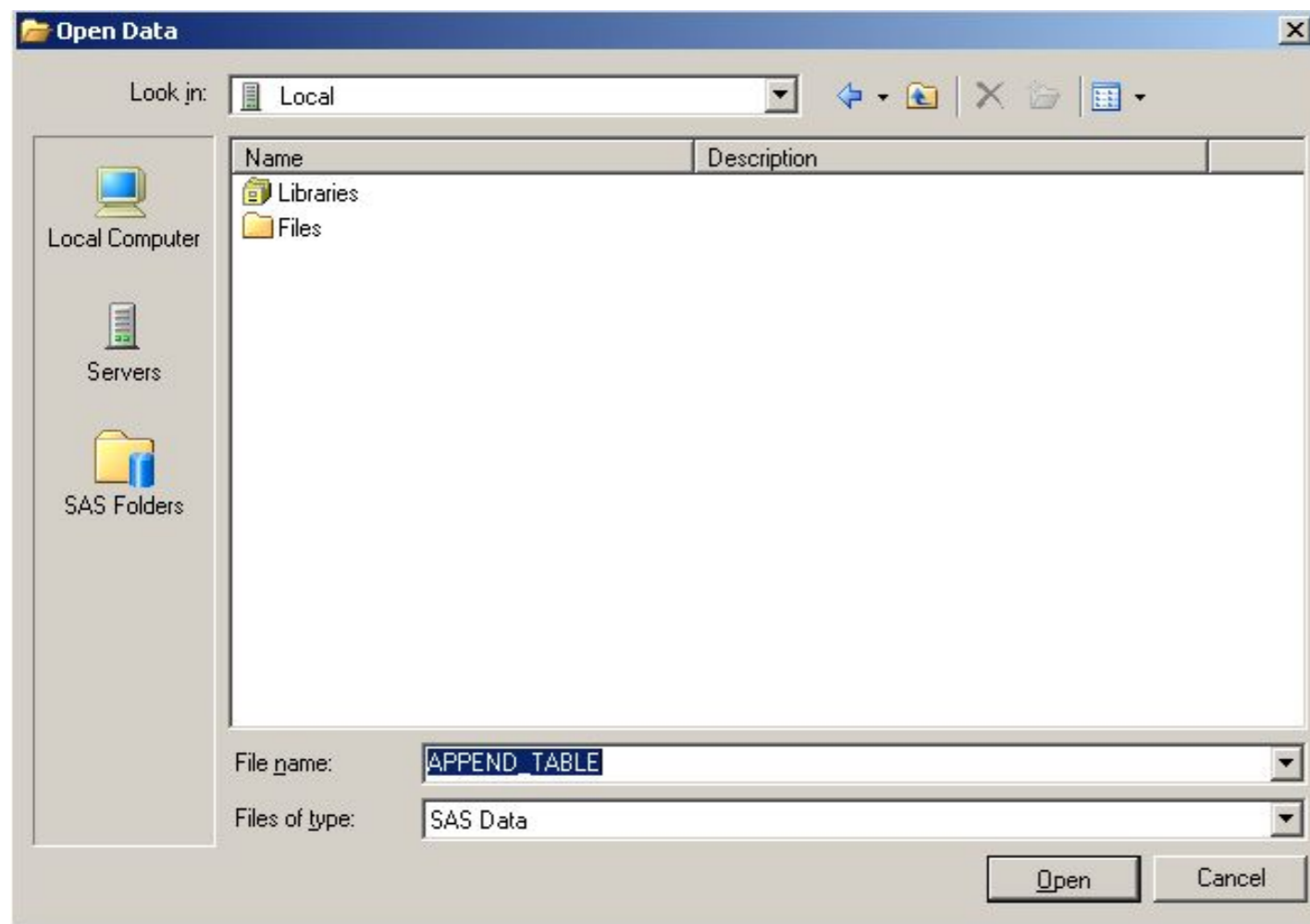
Открытие файлов из SAS папок

- SAS Папки позволяют получить доступ к данным и другим SAS метаданным зарегистрированными вашим SAS администратором



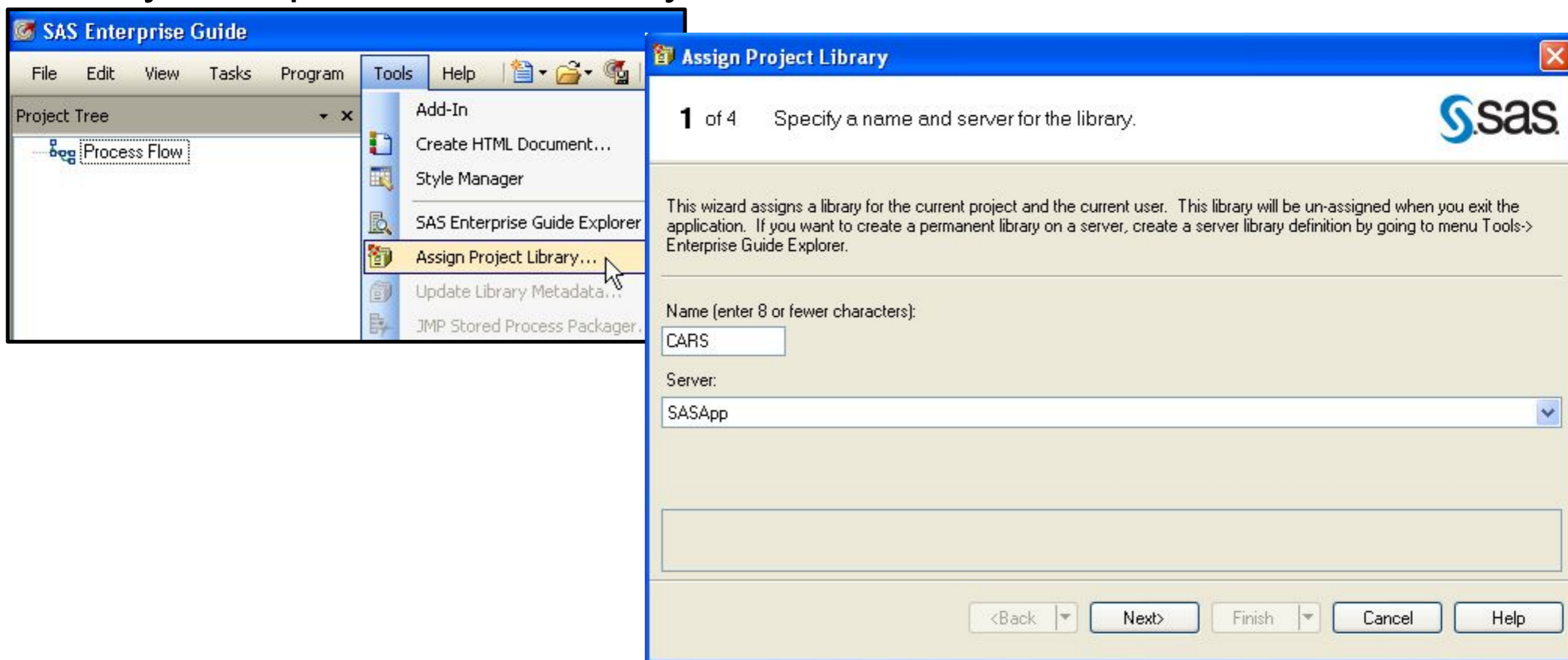
Доступ к удаленным данным

- Список серверов дает удобный путь для доступа ко всем вашим SAS файлам и таблицам с данными



Определение SAS библиотек

- SAS библиотеки обычно задаются SAS администратором. Однако в SAS EG ,есть функция Assign Project Library которая может быть использована для создания новых библиотек на основе SAS или прочих таблиц из внешних СУБД к которым нужно организовать доступ.



Доступ к удаленным данным

- После того как данные добавились в проект, они автоматически открываются в режиме просмотра (можно отключить в опциях)

The screenshot displays the SAS Enterprise Guide interface. The main window shows a data table with columns: Make, Model, Type, and C. The table contains 23 rows of car data. The 'Server List' window on the right shows a tree view of servers, including 'Local' and 'SASHELP' libraries. The 'SASHELP' library is expanded, showing various tables such as _CMPIDX_, ADOMSG, ADSMSG, AFMSG, AIR, ASSCMGR, BUY, BWEIGHT, CARS, CITIDAY, CITIMON, CITIQTR, CITIWK, CITYR, CLASS, CLASSFIT, CLNMSG, CNTAINER, and COLUMN. The 'CARS' table is highlighted in the list.

	Make	Model	Type	C
1	Acura	MDX	SUV	Asia
2	Acura	RSX Type S 2dr	Sedan	Asia
3	Acura	TSX 4dr	Sedan	Asia
4	Acura	TL 4dr	Sedan	Asia
5	Acura	3.5 RL 4dr	Sedan	Asia
6	Acura	3.5 RL w/Naviga...	Sedan	Asia
7	Acura	NSX coupe 2dr...	Sports	Asia
8	Audi	A4 1.8T 4dr	Sedan	Europe
9	Audi	A4 1.8T converti...	Sedan	Europe
10	Audi	A4 3.0 4dr	Sedan	Europe
11	Audi	A4 3.0 Quattro 4...	Sedan	Europe
12	Audi	A4 3.0 Quattro 4...	Sedan	Europe
13	Audi	A6 3.0 4dr	Sedan	Europe
14	Audi	A6 3.0 Quattro 4...	Sedan	Europe
15	Audi	A4 3.0 convertib...	Sedan	Europe
16	Audi	A4 3.0 Quattro c...	Sedan	Europe
17	Audi	A6 2.7 Turbo Qu...	Sedan	Europe
18	Audi	A6 4.2 Quattro 4...	Sedan	Europe
19	Audi	A8 L Quattro 4dr	Sedan	Europe
20	Audi	S4 Quattro 4dr	Sedan	Europe
21	Audi	RS 6 4dr	Sports	Europe
22	Audi	TT 1.8 convertib...	Sports	Europe
23	Audi	TT 1.8 Quattro 2	Sports	Europe

SAS таблицы - источники данных

SAS источники данных (таблицы) это прямоугольная таблица состоящая из строчек и колонок(столбцов)

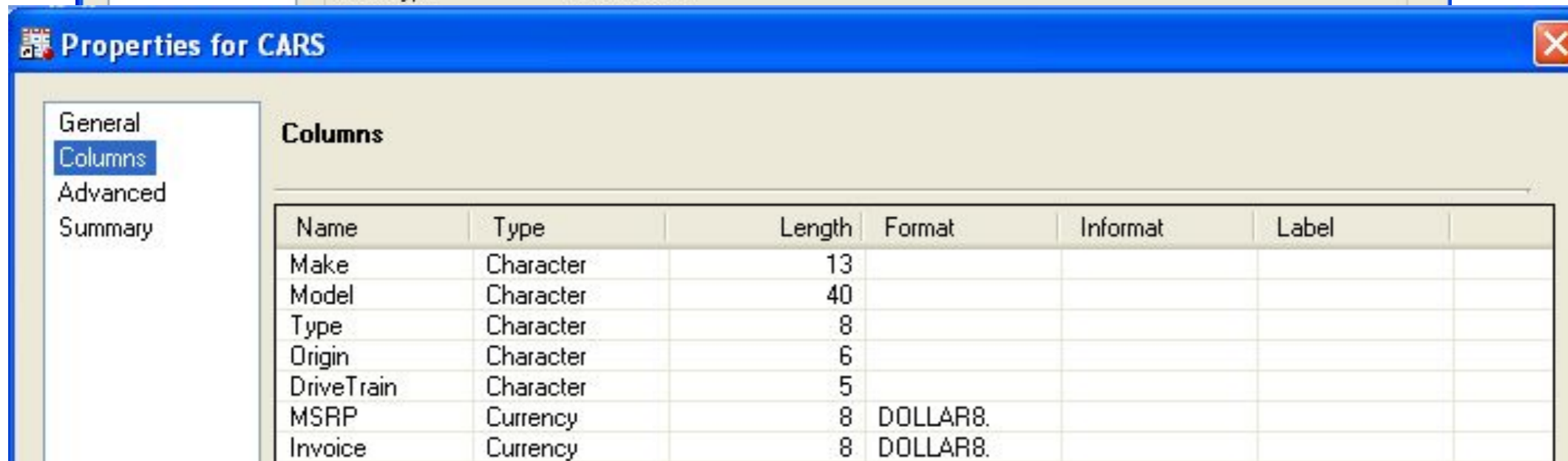
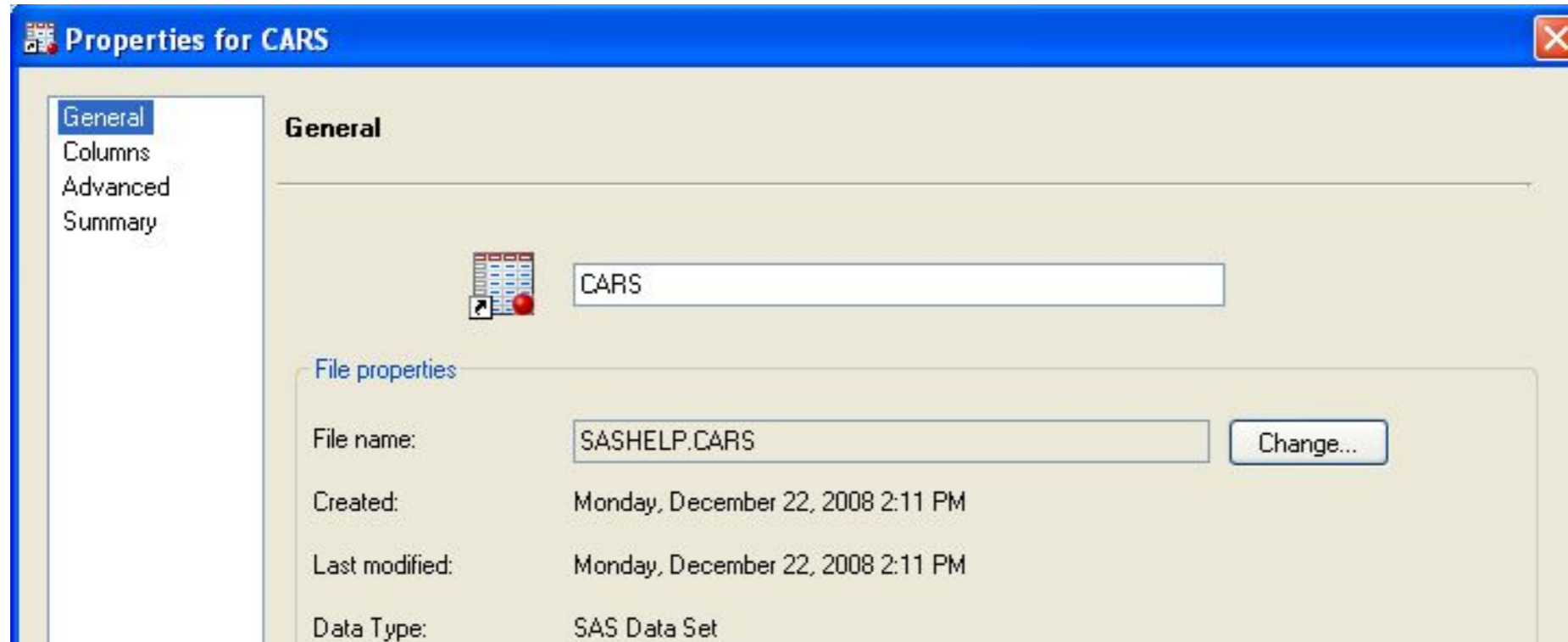
Origin	DriveTrain	MSRP	Invoice	EngineSize
Asia	All	\$36,945	\$33,337	3.5
Asia	Front	\$23,820	\$21,761	2
Asia	Front	\$26,990	\$24,647	2.4
Asia	Front	\$33,195	\$30,299	3.2
Asia	Front	\$43,755	\$39,014	3.5
Asia	Front	\$46,100	\$41,100	3.5
Asia	Rear	\$89,765	\$79,978	3.2
Europe	Front	\$25,940	\$23,508	1.8
Europe	Front	\$35,940	\$32,506	1.8
Europe	Front	\$31,840	\$28,846	3
Europe	All	\$33,430	\$30,366	3
Europe	All	\$34,480	\$31,388	3
Europe	Front	\$36,640	\$33,129	3

Строки
(Наблюдения)

Колонки
(Переменные)

SAS Свойства источника данных

- Просмотр свойств источника данных осуществляется нажатием правой кнопкой на нем, и позволяет посмотреть состав колонок и формат переменных.



Форматы

Formats

Categories:

- None
- Numeric
- Date**
- Time
- Date/Time
- Currency
- User Defined
- All

Formats:

- MINGUOw.
- MMDDYYBw.d
- MMDDYYCw.d
- MMDDYYDw.d
- MMDDYYNw.d
- MMDDYYPw.d
- MMDDYYSw.d
- MMDDYYw.d**

Attributes

Overall width: Min: 2 Max: 10

Decimal places: Min: 0 Max: 9

Description

date values

Example

Value: 14245 (01Jan1999)

Output:

0	1	/	0	1	/	1	9	9	9
---	---	---	---	---	---	---	---	---	---

OK Cancel

Formats

Categories:

- None
- Numeric
- Date
- Time
- Date/Time
- Currency**
- User Defined
- All

Formats:

- DOLLARw.d**
- DOLLARXw.d
- EURFRATSw.d
- EURFRBEFw.d
- EURFRCHFw.d
- EURFRCZKw.d
- EURFRDEMw.d
- EURFRDKKw.d

Attributes

Overall width: Min: 2 Max: 32

Decimal places: Min: 0 Max: 7

Description

dollar sign, commas and decimal point

Example

Value: 12345.1

Output:

1	2	3	4	5	.	1	0
---	---	---	---	---	---	---	---

OK Cancel

Создание форматов

- Помимо стандартных, возможно создавать собственные форматы с помощью задачи Create Format

Диалог импорта данных

- Для чтения текстовых файлов в SAS EG, необходимо указать структуру строчек и столбцов для чтения переменных в виде наблюдения и переменных.

```
100100001 1526 2 $3.50
100100001 3049 5 $8.75
100400004 8001 3 $5.25
100400004 8241 2 $3.50
100500005 4323 2 $3.50
100600006 7384 10 $17.50
100600006 6269 3 $5.25
100600006 1828 1 $1.75
100600006 1005 3 $5.25
```

ORDERINFORMATION (read-only)					
	Column1	Column2	ProdID	Quantity	SellPrice
1	1001	00001	1526	2	\$3.50
2	1001	00001	3049	5	\$8.75
3	1004	00004	8001	3	\$5.25
4	1004	00004	8241	2	\$3.50
5	1005	00005	4323	2	\$3.50
6	1006	00006	7384	10	\$17.50
7	1006	00006	6269	3	\$5.25
8	1006	00006	1828	1	\$1.75
9	1006	00006	1005	3	\$5.25

Диалог импорта данных

- Внешний вид диалог импорта, два варианта импорт файла с разделителями и фиксированной ширины.

Import Data from adult.csv

2 of 4 Select Data Source

sas

Text format

Delimited fields

Space

Fixed columns

File contains field names on record number: 1

Data records start at record number: 2

Limit the number of records read to:

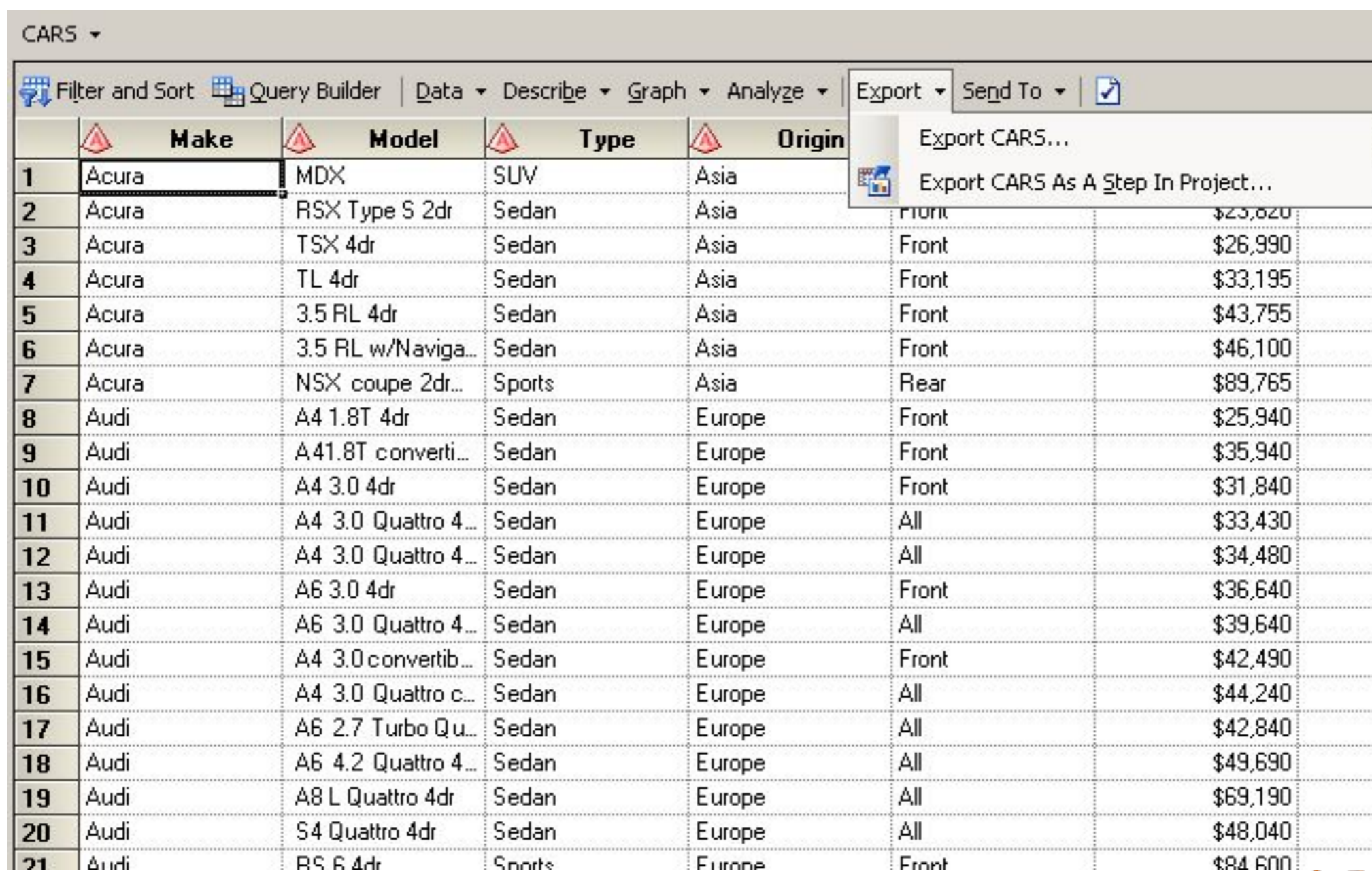
Rename columns to comply with SAS naming conventions.

```
Age Workclass Fnlwgt Education Education-num Martial-status Occupation Relations
25 Private 226802 11th 7 Never-married Machine-op-inspct Own-child Black Male 0
38 Private 89814 HS-grad 9 Married-civ-spouse Farming-fishing Husband White Male
28 Local-gov 336951 Assoc-acdm 12 Married-civ-spouse Protective-serv Husband Whi
44 Private 160323 Some-college 10 Married-civ-spouse Machine-op-inspct Husband B
18 ? 103497 Some-college 10 Never-married ? Own-child White Female 0 0 30 United
34 Private 198693 10th 6 Never-married Other-service Not-in-family White Male 0
29 ? 227026 HS-grad 9 Never-married ? Unmarried Black Male 0 0 40 United-States
63 Self-emp-not-inc 104626 Prof-school 15 Married-civ-spouse Prof-specialty Husb
24 Private 369667 Some-college 10 Never-married Other-service Unmarried White Fe
55 Private 104996 7th-8th 4 Married-civ-spouse Craft-repair Husband White Male 0
65 Private 184454 HS-grad 9 Married-civ-spouse Machine-op-inspct Husband White M
36 Federal-gov 212465 Bachelors 13 Married-civ-spouse Adm-clerical Husband White
26 Private 82091 HS-grad 9 Never-married Adm-clerical Not-in-family White Female
58 ? 299831 HS-grad 9 Married-civ-spouse ? Husband White Male 0 0 35 United-Stat
48 Private 279724 HS-grad 9 Married-civ-spouse Machine-op-inspct Husband White M
43 Private 346189 Masters 14 Married-civ-spouse Exec-managerial Husband White Ma
```

<Back Next> Finish Cancel Help

Экспорт данных в файл

- При необходимости результаты работы аналитического узла можно выгрузить во внешний файл, либо в SAS таблицу.



CARS ▾

Filter and Sort Query Builder | Data ▾ Describe ▾ Graph ▾ Analyze ▾ Export ▾ Send To ▾

	Make	Model	Type	Origin		
1	Acura	MDX	SUV	Asia		
2	Acura	RSX Type S 2dr	Sedan	Asia	Front	\$25,020
3	Acura	TSX 4dr	Sedan	Asia	Front	\$26,990
4	Acura	TL 4dr	Sedan	Asia	Front	\$33,195
5	Acura	3.5 RL 4dr	Sedan	Asia	Front	\$43,755
6	Acura	3.5 RL w/Naviga...	Sedan	Asia	Front	\$46,100
7	Acura	NSX coupe 2dr...	Sports	Asia	Rear	\$89,765
8	Audi	A4 1.8T 4dr	Sedan	Europe	Front	\$25,940
9	Audi	A4 1.8T converti...	Sedan	Europe	Front	\$35,940
10	Audi	A4 3.0 4dr	Sedan	Europe	Front	\$31,840
11	Audi	A4 3.0 Quattro 4...	Sedan	Europe	All	\$33,430
12	Audi	A4 3.0 Quattro 4...	Sedan	Europe	All	\$34,480
13	Audi	A6 3.0 4dr	Sedan	Europe	Front	\$36,640
14	Audi	A6 3.0 Quattro 4...	Sedan	Europe	All	\$39,640
15	Audi	A4 3.0 convertib...	Sedan	Europe	Front	\$42,490
16	Audi	A4 3.0 Quattro c...	Sedan	Europe	All	\$44,240
17	Audi	A6 2.7 Turbo Qu...	Sedan	Europe	All	\$42,840
18	Audi	A6 4.2 Quattro 4...	Sedan	Europe	All	\$49,690
19	Audi	A8 L Quattro 4dr	Sedan	Europe	All	\$69,190
20	Audi	S4 Quattro 4dr	Sedan	Europe	All	\$48,040
21	Audi	RS 6 4dr	Sports	Europe	Front	\$84,600

Изменение формата и типов колонок

- Настройки форматов колонок, можно настроить фильтром или запросом.

The screenshot displays the SAS Query Builder1 interface for the Local:SASHELP.CARS dataset. The main window shows a list of columns with their respective input and summary fields. A 'Formats' dialog box is open, showing the 'Numeric' category selected. The 'Formats' list includes various SAS date and time formats, with 'BESTw.d' selected. The 'Attributes' section shows 'Overall width' set to 12 and 'Decimal places' set to 0. The 'Description' field contains the text 'SAS System chooses best notation'. The 'Example' section shows a value of 123.1 and its corresponding output format: '123.1'.

Column Name	Input	Summary	Format
Make	t1.Make		
Model	t1.Model		
Type	t1.Type		
Origin	t1.Origin		
DriveTrain	t1.DriveTrain		
MSRP	t1.MSRP		
Invoice	t1.Invoice		
EngineSize	t1.EngineSize		
Cylinders	t1.Cylinders		
Horsepower	t1.Horsepower		
MPG_City	t1.MPG_City		
MPG_Highway	t1.MPG_Highway		

Properties for Cylinders

Alias: Cylinder_per_car
Label: Cylinders in Car
Format: None
Summary: None Length (in t...
Expression:
Input: t1.Cylinders

Formats

Categories: None, **Numeric**, Date, Time, Date/Time, Currency, User Defined, All

Formats: B8601DAw., B8601DNw., B8601DTw.d, B8601DZw.d, B8601TMw.d, B8601TZw.d, BESTDw.d, **BESTw.d**

Attributes: Overall width: 12 Min: 1 Max: 32
Decimal places: 0 Min: 0 Max: 11

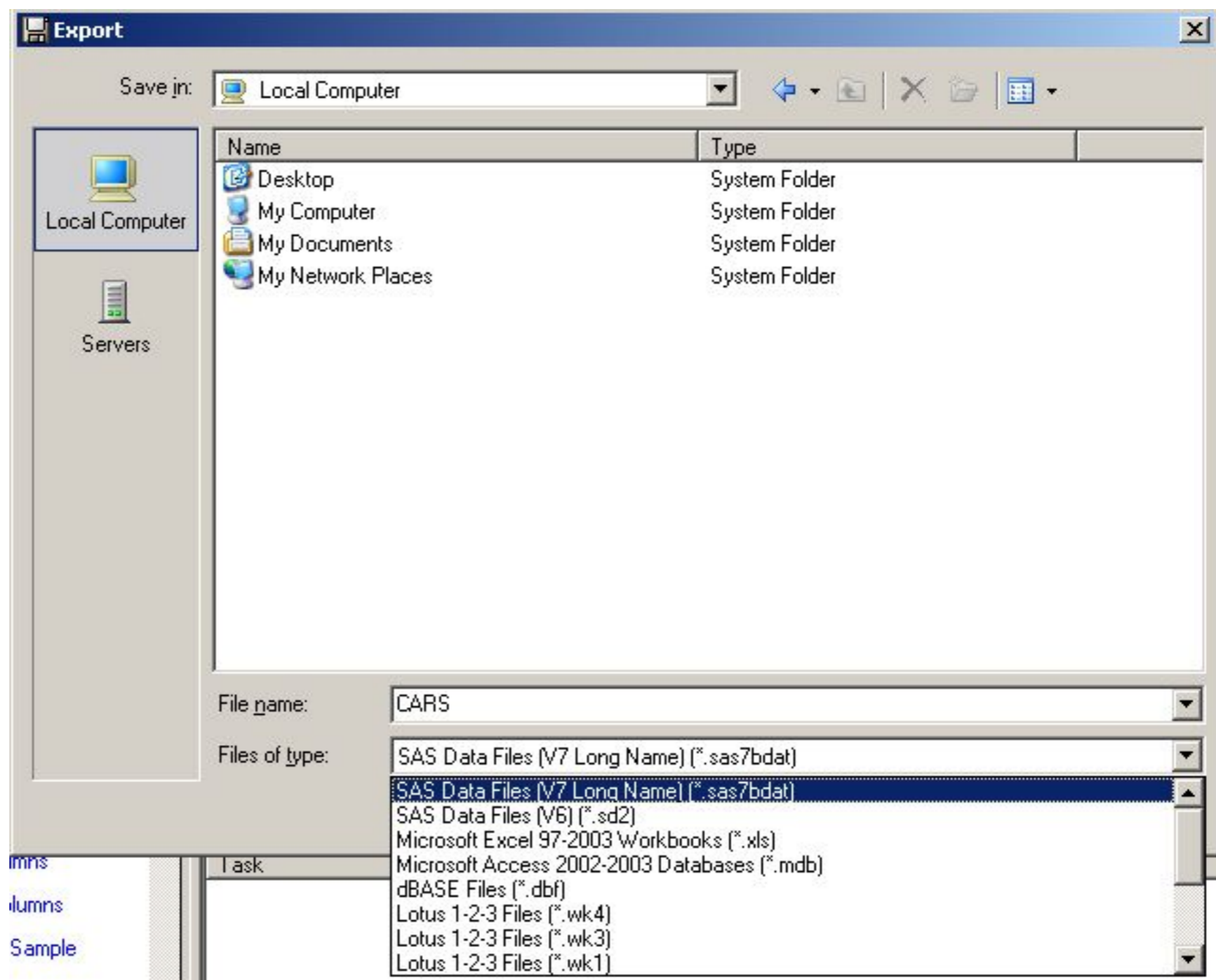
Description: SAS System chooses best notation

Example: Value: 123.1
Output: 123.1

Настройка формата для экспорта в файл

Затем при экспорте можно указать положение на диске и формат экспортируемых результатов:

- SAS таблицы
- Excel таблицы
- Access файлы
- DBF файлы
- Csv файлы
- Текстовый файлы.



Раздел 3: Работа с запросами

3.1. Введение в запросы

3.2. Фильтрация и сортировка

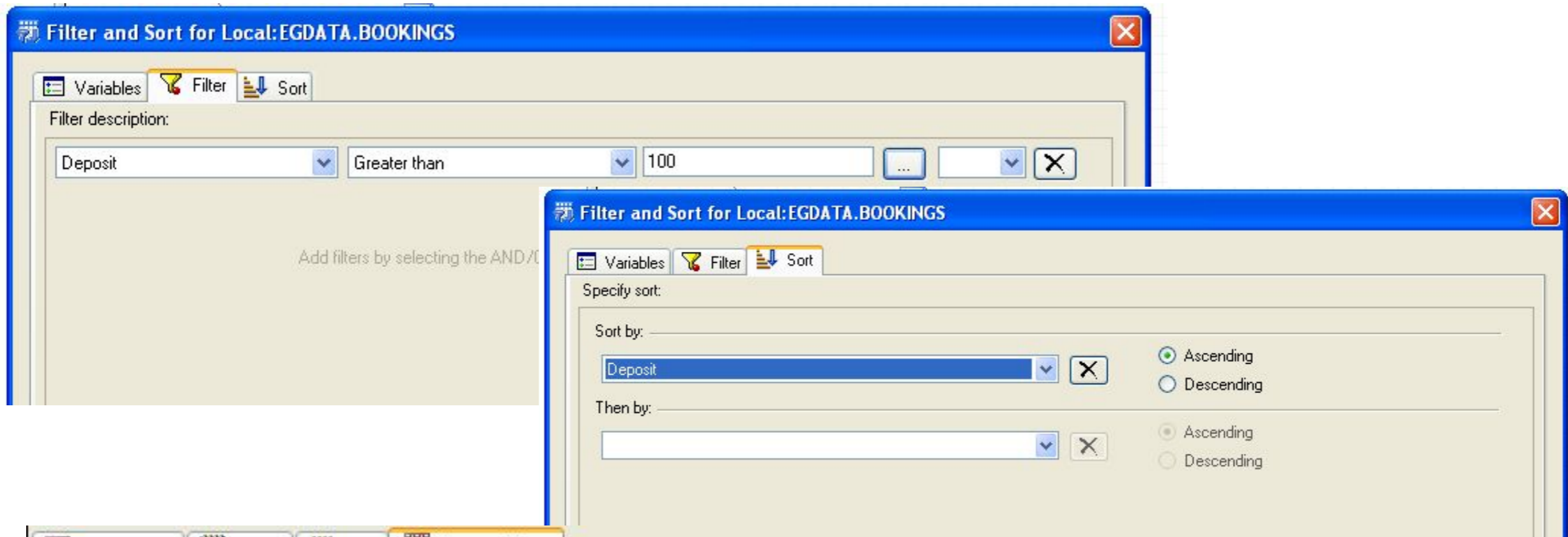
3.3. Создание колонок с помощью вычисляемых выражений

3.4. Группировка и суммаризация в запросе

3.5. Объединение таблиц

Фильтры, сортировка и редактор запросов

- Фильтры и сортировка можно использовать для создания новых таблиц, на основе выбранного критерия фильтрации или сортировки



	Office	CustomerID	Tour	Travelers	Deposit	Deposit_Date
1	Portland	HI15	SH43	4	150	31JUL2006:12:0...
2	Portland	DI33	SH43	4	150	23JUL2006:12:0...
3	Portland	DE31	FJ12	3	175	25JUL2006:12:0...
4	Portland	SL34	FJ12	4	200	19JUL2006:12:0...
5	Portland	GI08	SH43	8	300	31JUL2006:12:0...
6	Portland	SL28	SH43	10	425	05JUL2006:12:0...

Фильтры и сортировка

- Выбираем переменные для фильтрации.

Filter and Sort for Local:EGDATA.BOOKINGS

Variables Filter Sort

Available (0):

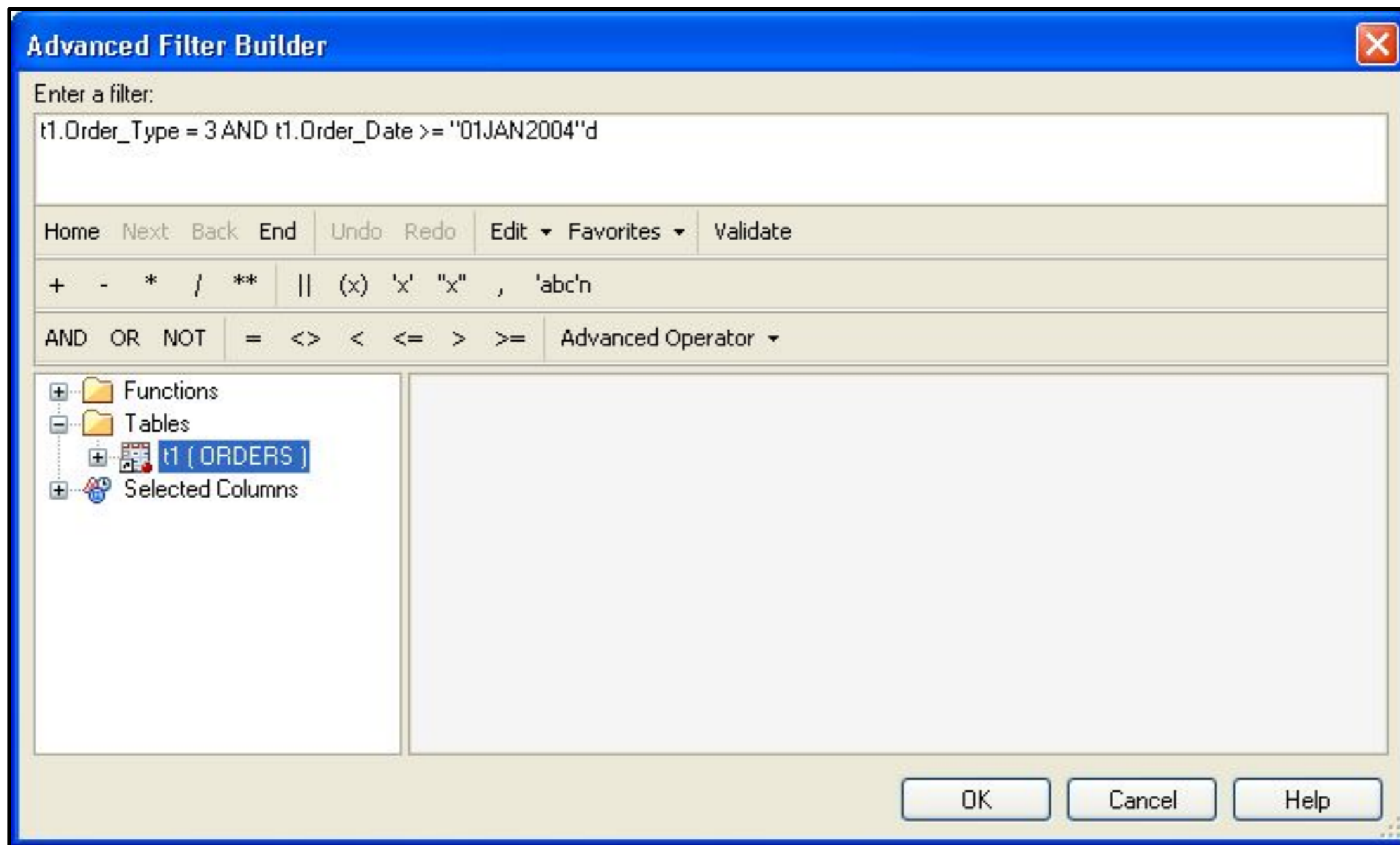
Name

Selected (6 of 6):

Name	Type	Label
Office	Char	Office
CustomerID	Char	CustomerID
Tour	Char	Tour
Travelers	Num	Travelers
Deposit	Num	Deposit
Deposit_Date	Date	Deposit_Date

Расширенный редактор фильтра.

- Расширенный режим фильтра позволяет создавать фильтры на основе SAS функций, формул и логических выражений



Фильтр и сортировка: Вкладка сортировки

- Вы можете сортировать сразу по нескольким переменным и выбирать требуемый порядок сортировки.

The screenshot shows the 'Filter and Sort for Local:SASUSER.ORDERS' dialog box with the 'Sort' tab selected. The dialog has three tabs: 'Variables', 'Filter', and 'Sort'. The 'Specify sort:' section contains three rows for sorting criteria. The first row is 'Sort by: Order_Date' with 'Ascending' selected. The second row is 'Then by: Profit' with 'Descending' selected. The third row is 'Then by:' with an empty dropdown and 'Ascending' selected. At the bottom, there are buttons for 'Show Preview', 'Validate', 'Clear All', 'OK', 'Cancel', and 'Help'. A checkbox for 'Display labels instead of variable names' is unchecked.

Filter and Sort for Local:SASUSER.ORDERS

Variables Filter Sort

Specify sort:

Sort by: Order_Date Ascending Descending

Then by: Profit Ascending Descending

Then by: Ascending Descending

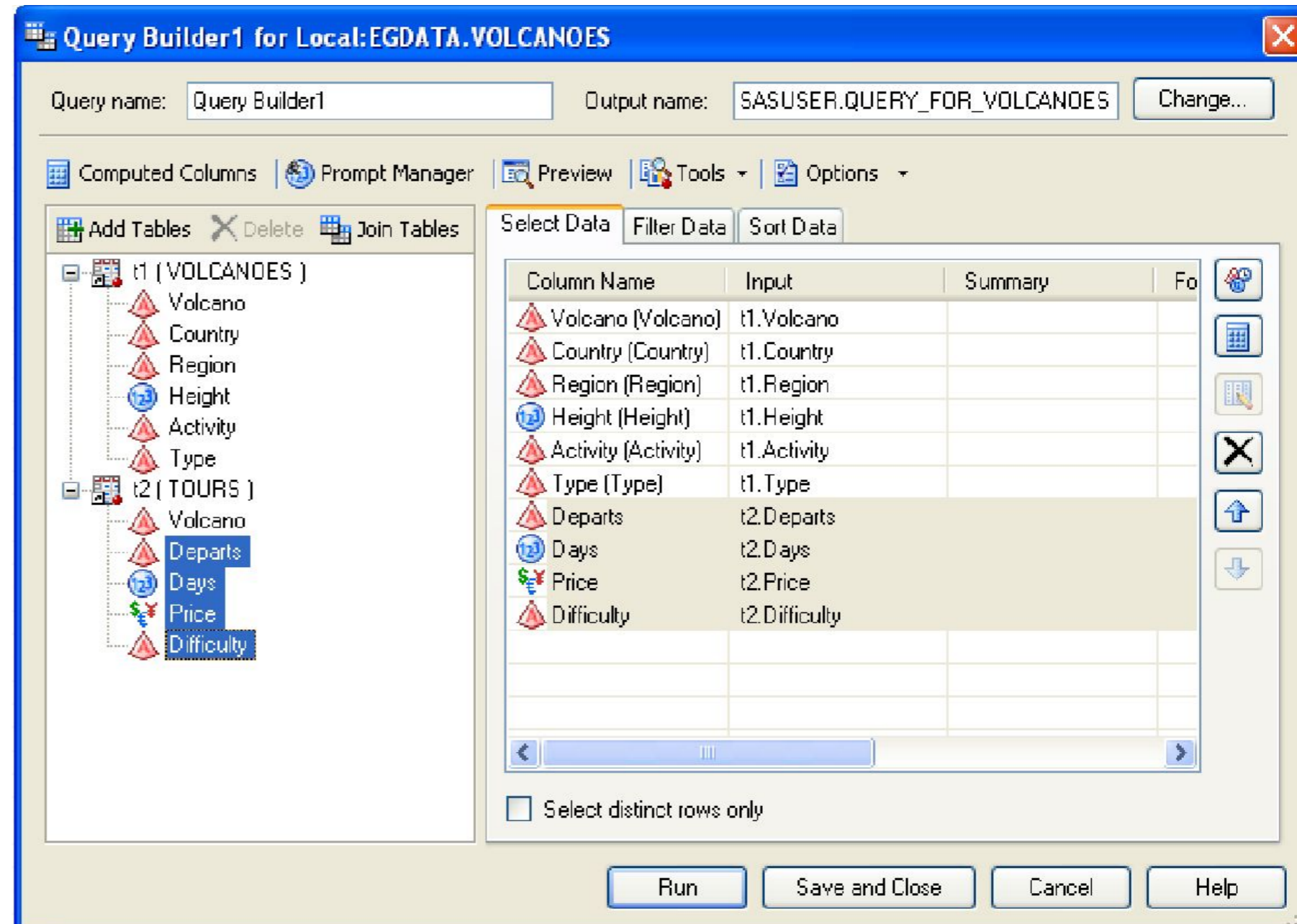
Display labels instead of variable names

Clear All

Show Preview Validate OK Cancel Help

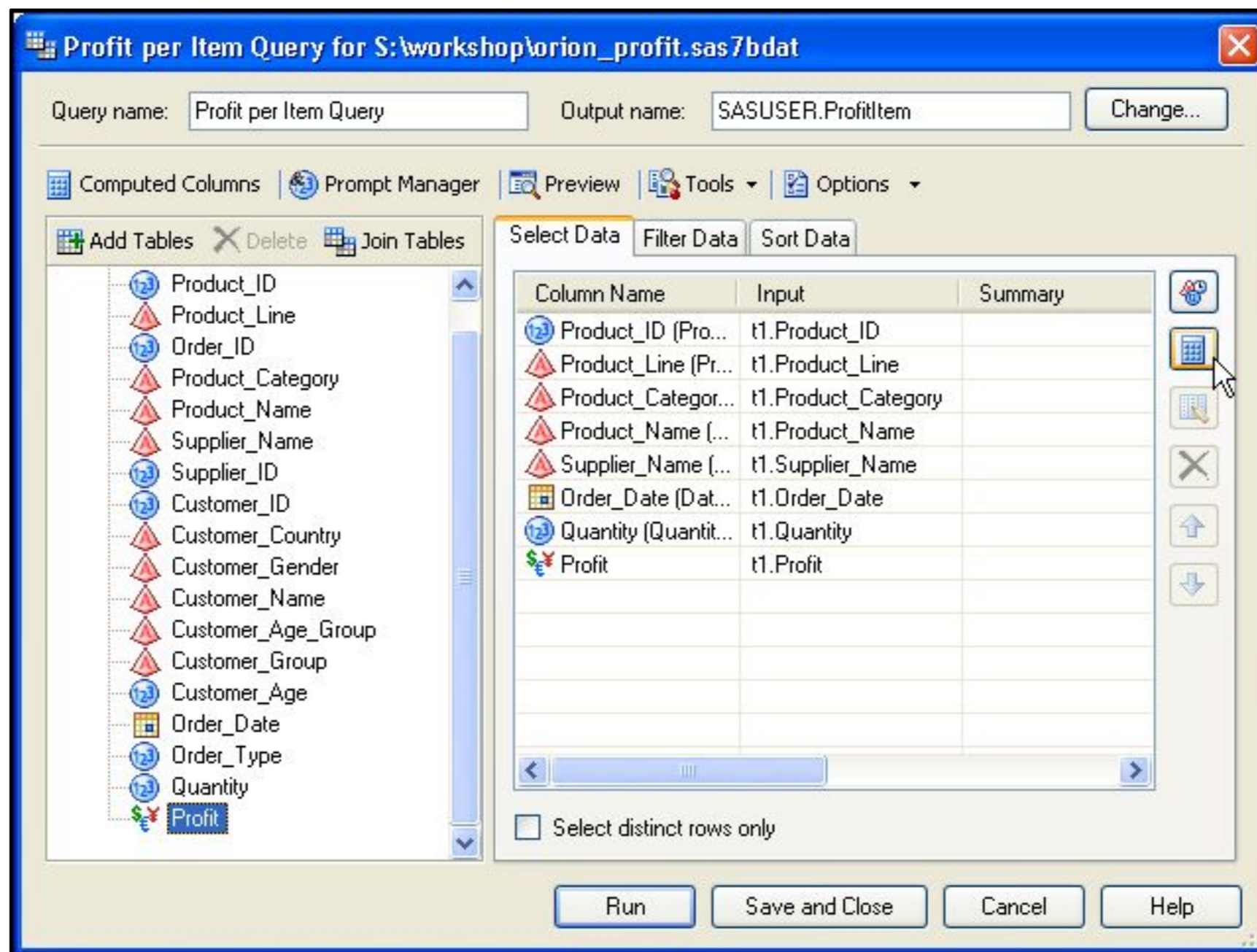
Редактор запроса

- Редактор запроса позволяет:
- Модифицировать свойства колонок
- Группировать и суммаризировать данные
- изменять форматы
- Выбирать разные записи (distinct)
- Объединять таблицы



Вычисляемые колонки (в редакторе запросов)

- С помощью редактора запросов можно создавать вычисляемые колонки: суммаризации, формулы и перекодировку. Для этого надо нажать кнопку Вычисляемые колонки.



Настройка новой вычисляемой колонки

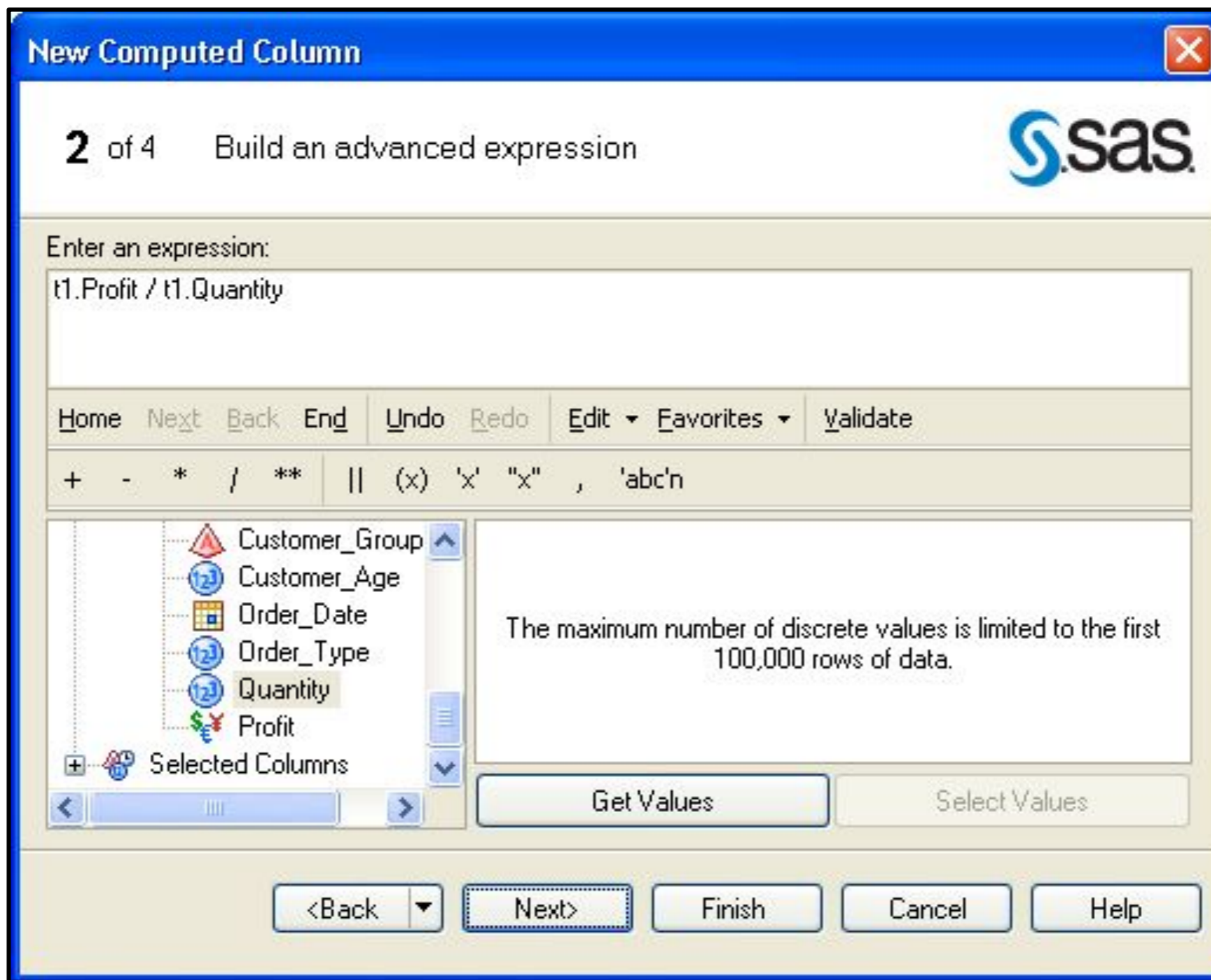
- Настройка вычисляемых колонок включает выбор типа выражения, задания формулы, агрегации или рекодировки, и выбор названия с форматом для новой колонки.

The image displays three overlapping screenshots of the 'New Computed Column' dialog box in SAS, illustrating the configuration steps:

- Step 1: Select a type**
 - Options: Summarized column, Recoded column, Advanced expression, From an existing computed column.
 - Buttons: <Back, Next>
- Step 2: Build an advanced expression**
 - Enter an expression: t1.Deposit
 - Tools: Home, Next, Back, End, Undo, Redo, Edit.
 - Operators: +, -, *, /, **, ||, (x), 'x', "x", ,.
 - Tree view: CustomerID, Tour, Travelers, Deposit, Deposit_Date, Selected Columns, Computed Columns.
 - Buttons: <Back, Next>
- Step 3: Modify additional options**
 - Column: Log_deposit
 - Alias: Log_deposit|
 - Summary: NONE (dropdown), Length (in bytes):
 - Expression: LOG(t1.Deposit)
 - Column type: Numeric, Character.
 - Format: None (dropdown), Change...
 - Buttons: <Back, Next>, Finish, Cancel, Help.

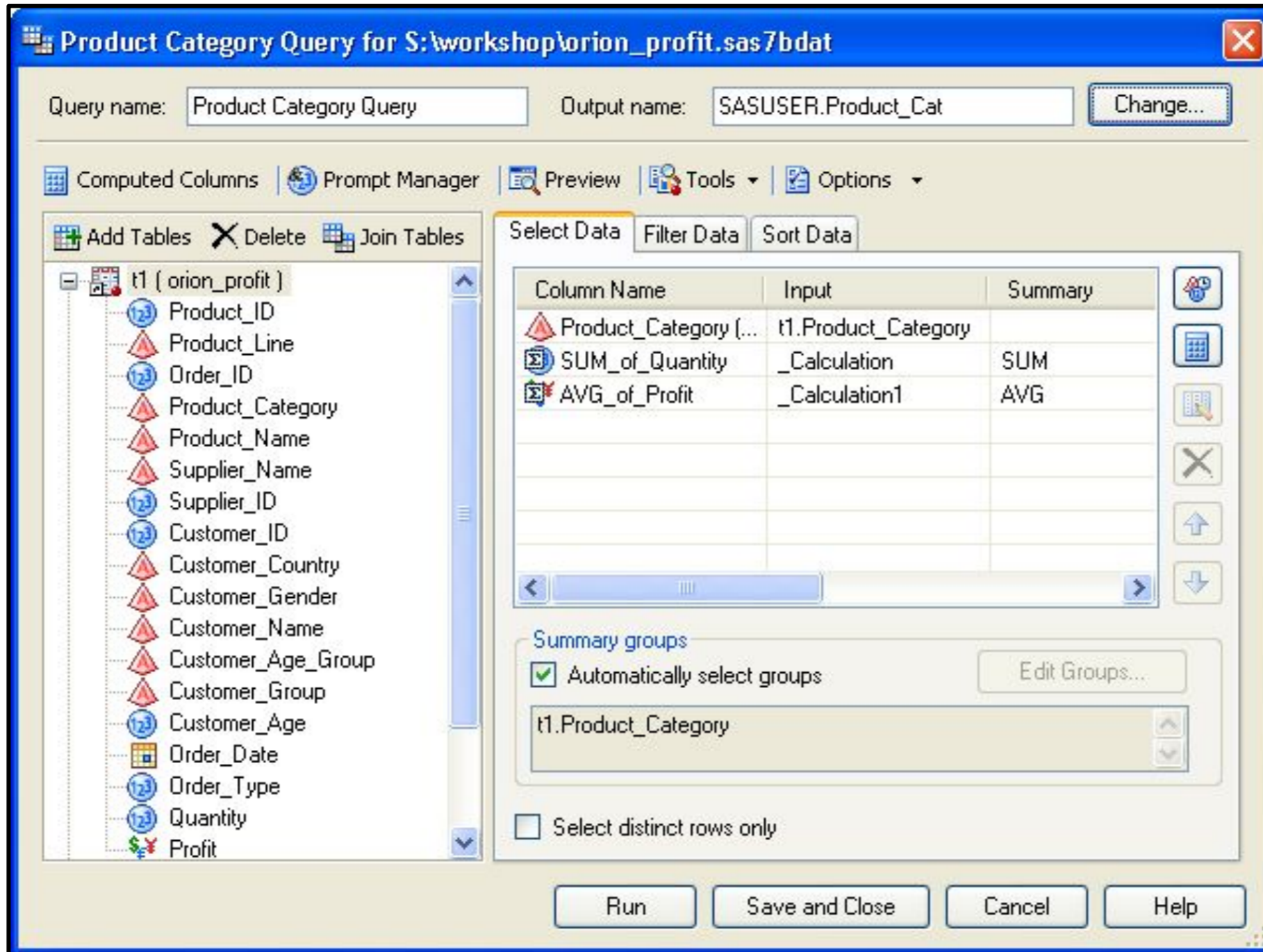
Редактор формульных выражений

- Редактор формул позволяет создавать простые и сложные выражения на основе встроенных в SAS функций и логических выражений.



Группировка данных

- В случае задания агрегатных функций, автоматически выбираются правила группировки на закладке выбора переменных



Фильтрация данных

- На основе суммаризированных показателей можно строить фильтры.

The screenshot displays the SAS Query Builder interface for a query named "Product Category Query". The window title is "Product Category Query for S:\workshop\orion_profit.sas7bdat". The query name is "Product Category Query" and the output name is "SASUSER.Product_Cat".

The interface includes a list of columns on the left, such as Product_Category, Product_Name, Supplier_Name, Supplier_ID, Customer_ID, Customer_Country, Customer_Gender, Customer_Name, Customer_Age_Group, Customer_Group, Customer_Age, Order_Date, Order_Type, Quantity, Profit, and Computed Columns (Calculation, _Calculation1).

The "Filter Data" tab is active, showing two filter sections:

- Filter the raw data:** A filter is applied using the "Where" clause: `t1.Order_Type = 3`.
- Filter the summarized data:** A filter is applied using the "Having" clause: `(CALCULATED SUM_of_Quantity) > 5`.

The interface also includes buttons for "Run", "Save and Close", "Cancel", and "Help".

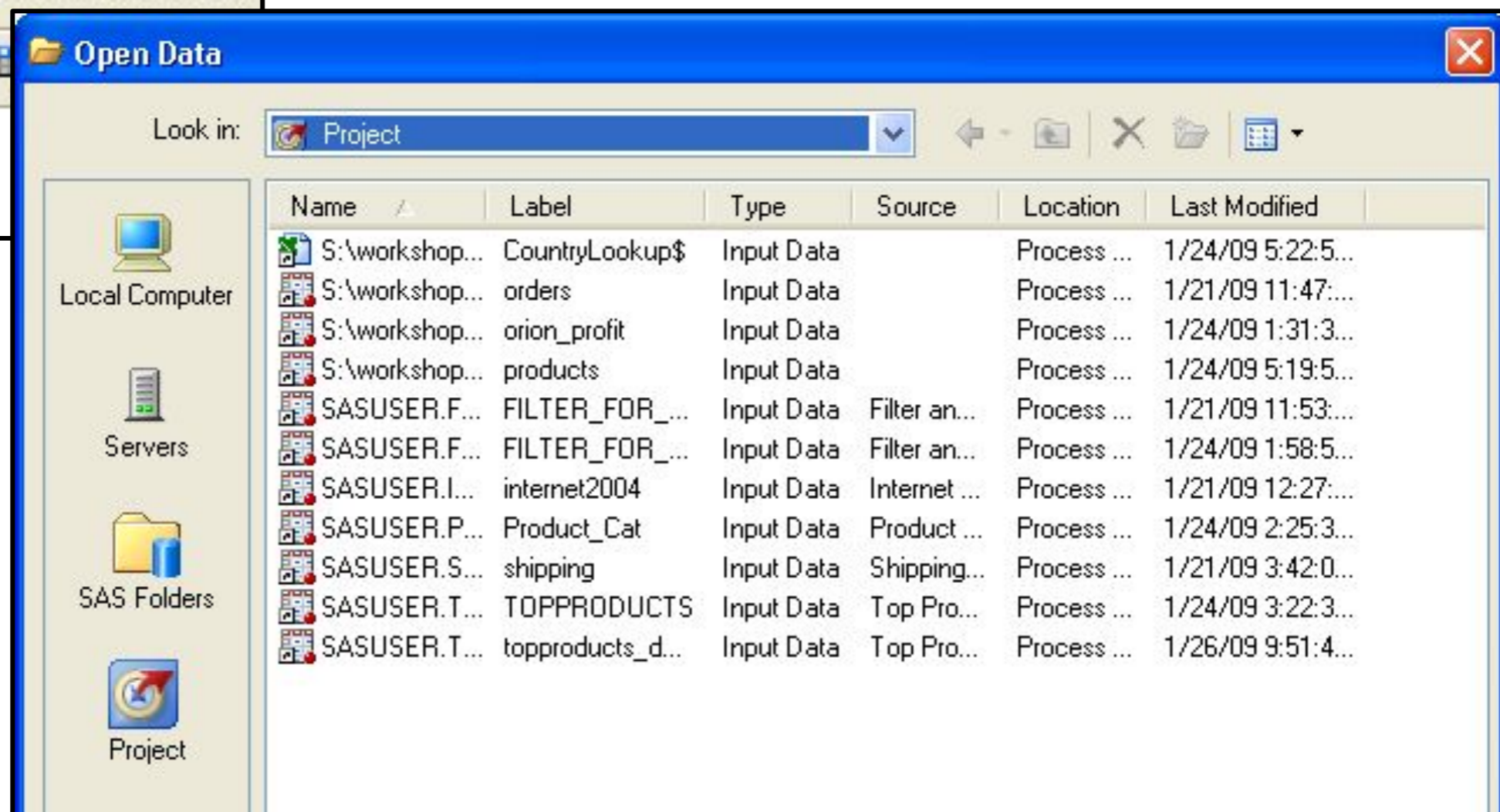
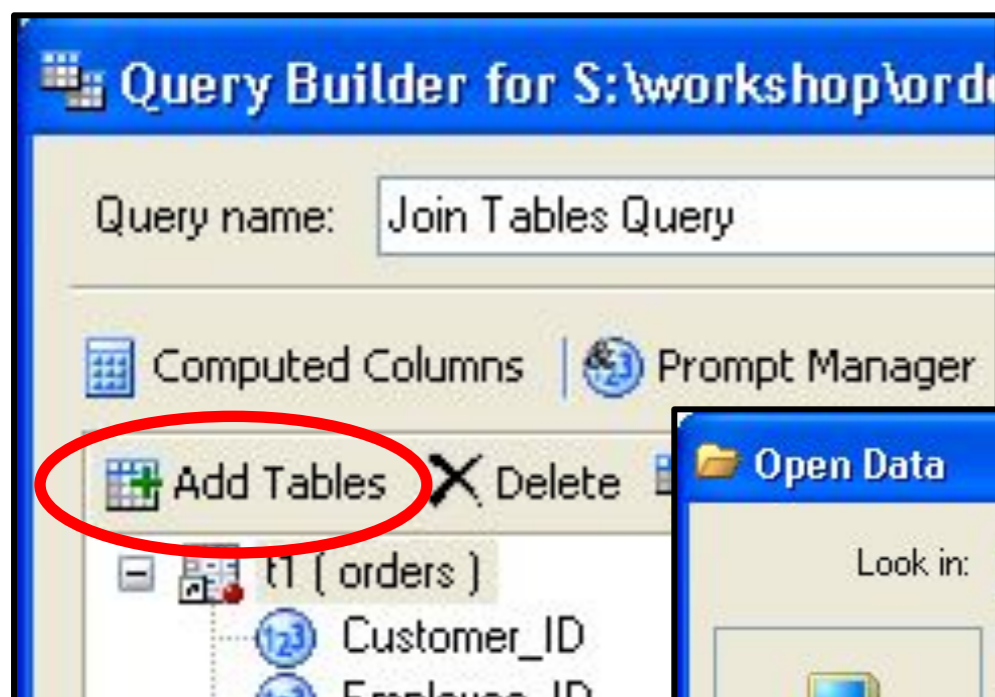
Объединение таблиц

- Прямо в редакторе запросов можно настраивать порядок и тип объединений для таблиц

The screenshot displays a database query editor window titled "Tables and Joins". It shows two tables, t1 (VOLCANOES) and t2 (TOURS), connected by an inner join symbol. The t1 table has columns: Volcano, Country, Region, Height, Activity, and Type. The t2 table has columns: Volcano, Departs, Days, Price, and Difficulty. A "Join Properties" dialog box is open, showing the "Join type" as "Matching rows only given a condition (Inner Join)". The "Condition" section shows "t1.Volcano" joined to "t2.Volcano" with an equals sign. The dialog also includes a "Filter to include in the 'join tables on' clause" field, "Clear..." and "Edit..." buttons, and "OK", "Cancel", and "Help" buttons at the bottom.

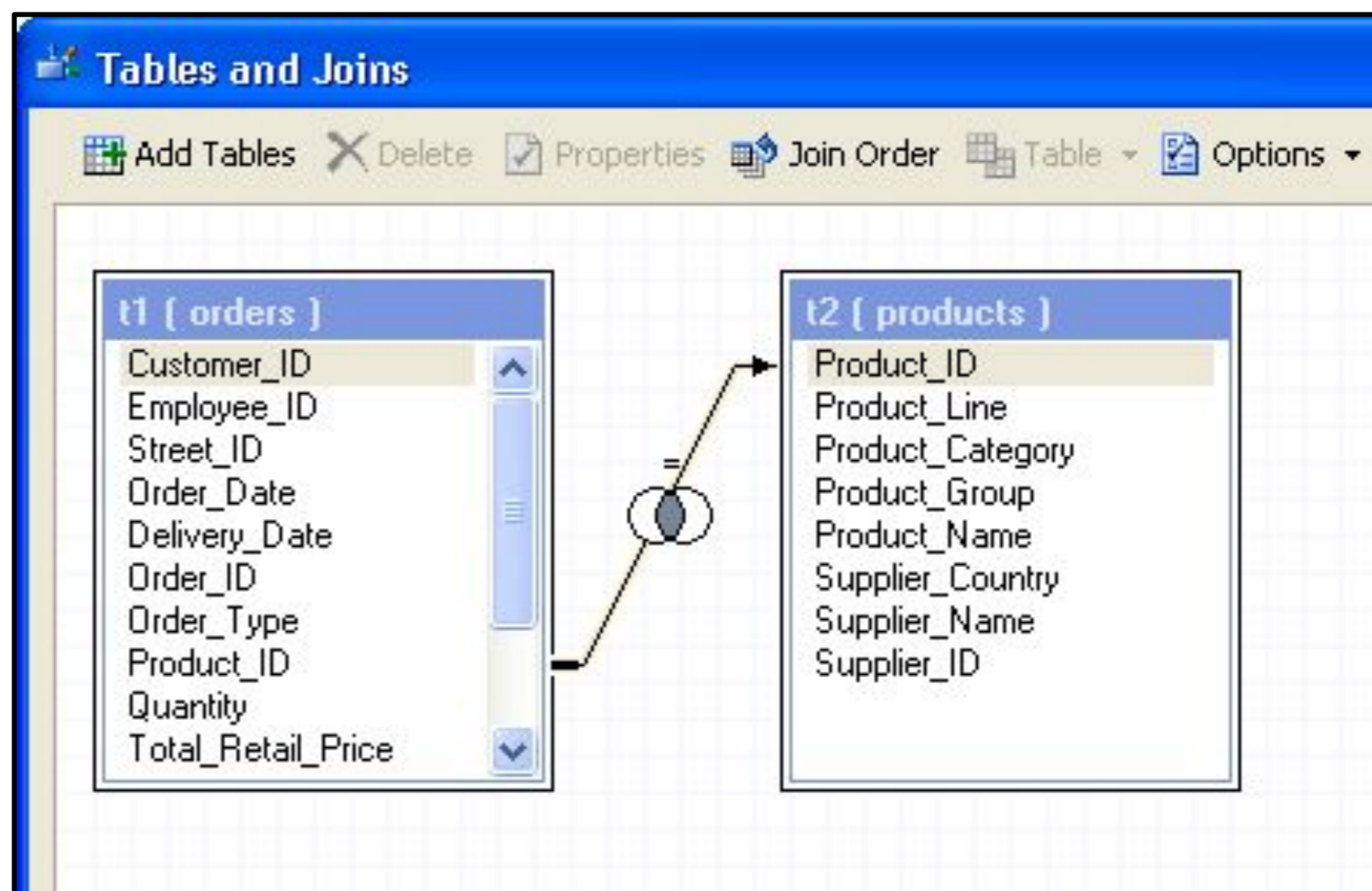
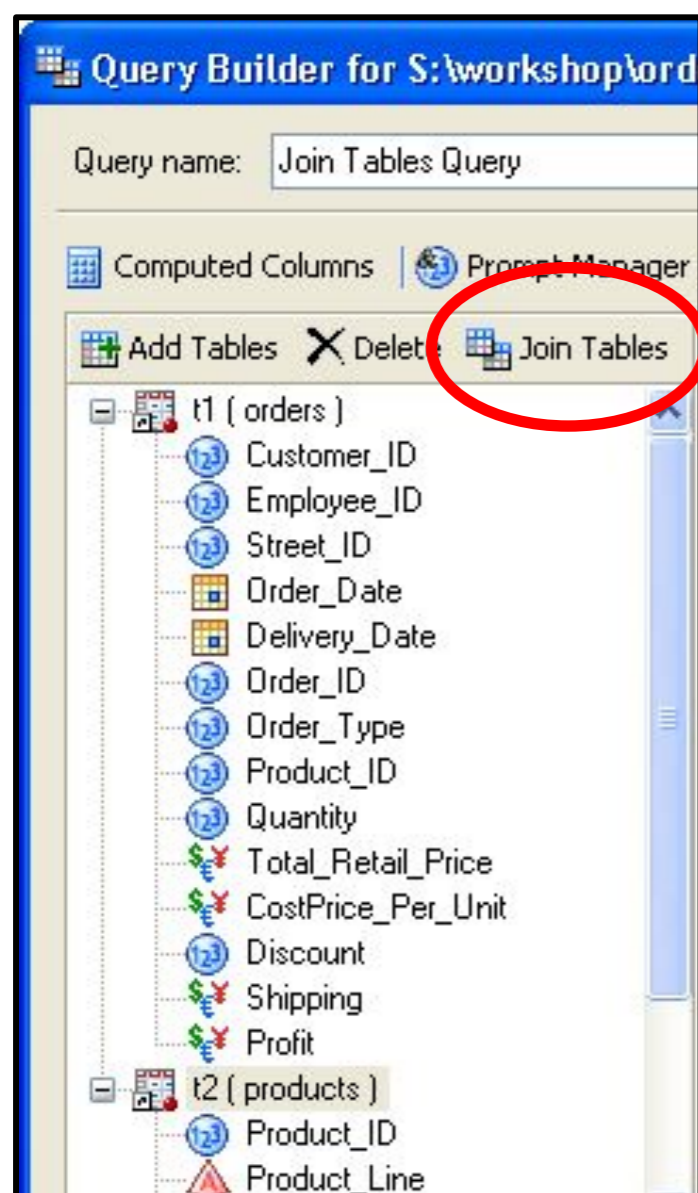
Добавление таблиц к объединению

- Нажмите кнопку Добавить таблицы, чтобы включить в запрос нужное количество источников данных



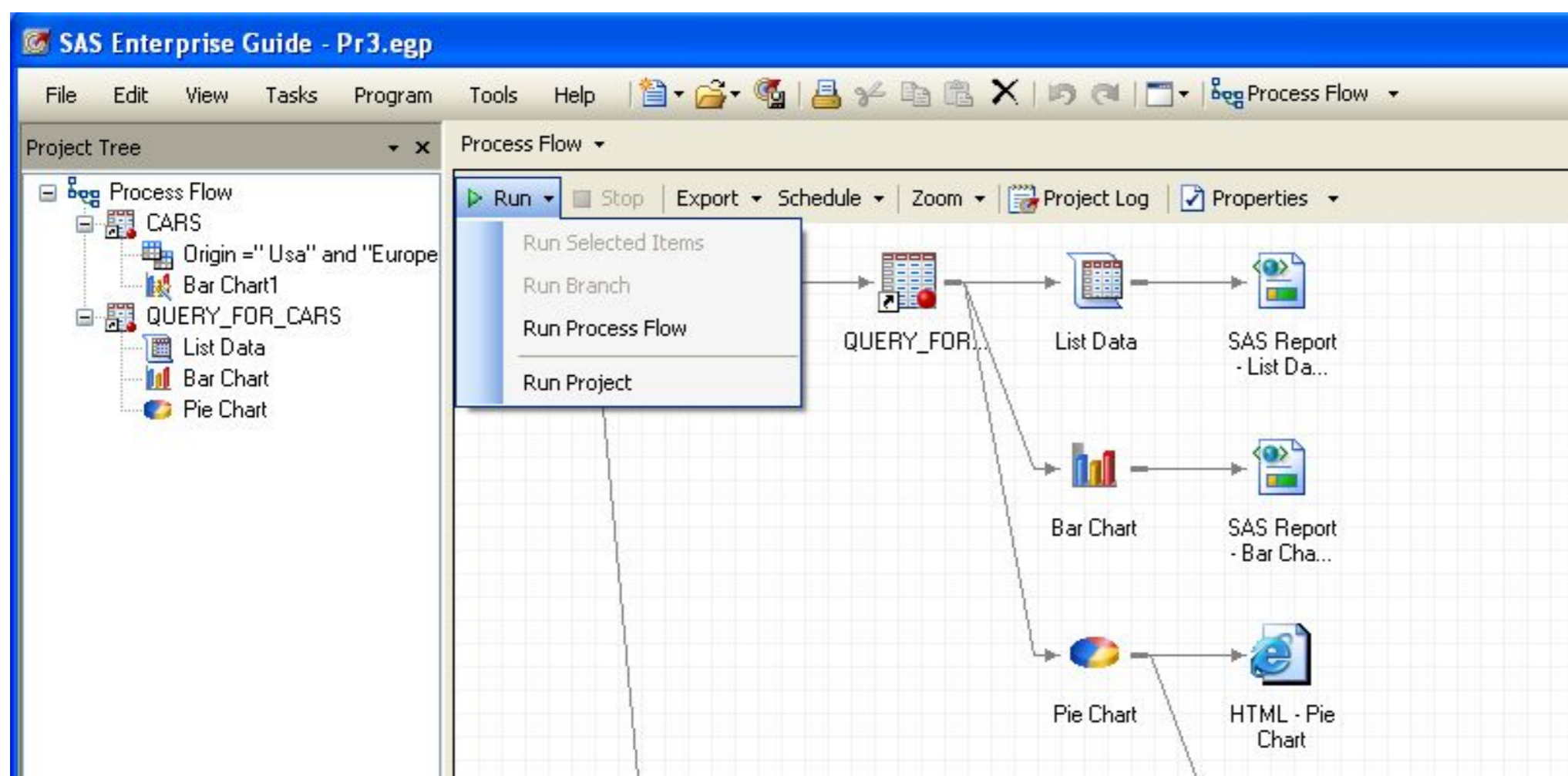
Таблицы и окно объединения

- Нажмите кнопку Join Tables, и укажите требуемый порядок и тип объединения



Обновление результатов

- Нажатие кнопки **Run** позволяет выполнять обновления в следующих режимах:
 - Для выделенных элементов (узлов) проекта
 - Для выделенных и всех последующих узлов в текущей диаграмме проекта
 - Для всего проекта целиком



Раздел 4. Работа с аналитическими задачами

4.1. Введение в аналитические задачи

4.2. Транспонирование

4.3. Создание отчета, частотного анализа

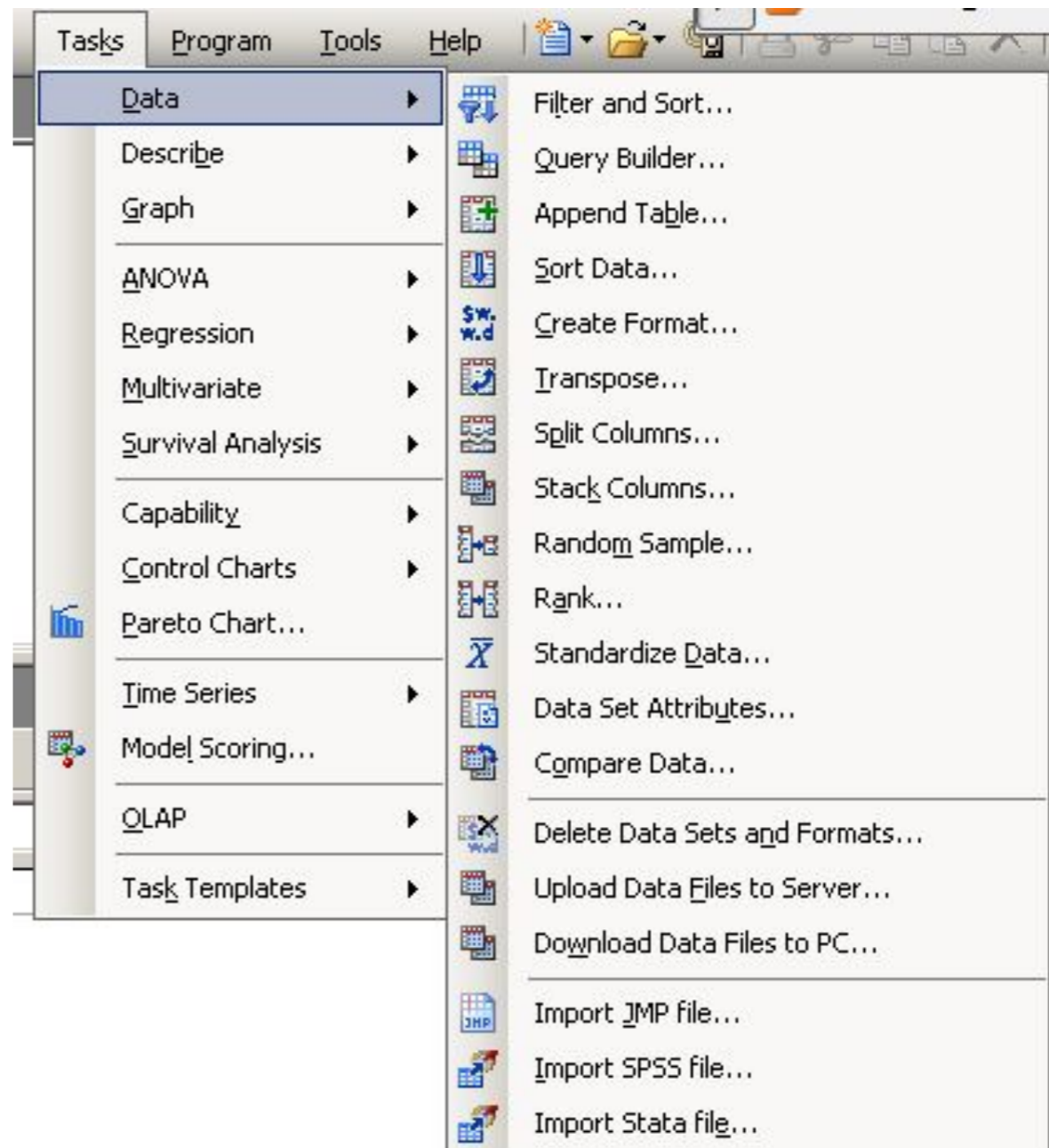
4.4. Создание отчетов в форматах HTML, PDF и RTF

4.5. Сводная статистика и таблица

Задачи группы Data - манипуляций с данными

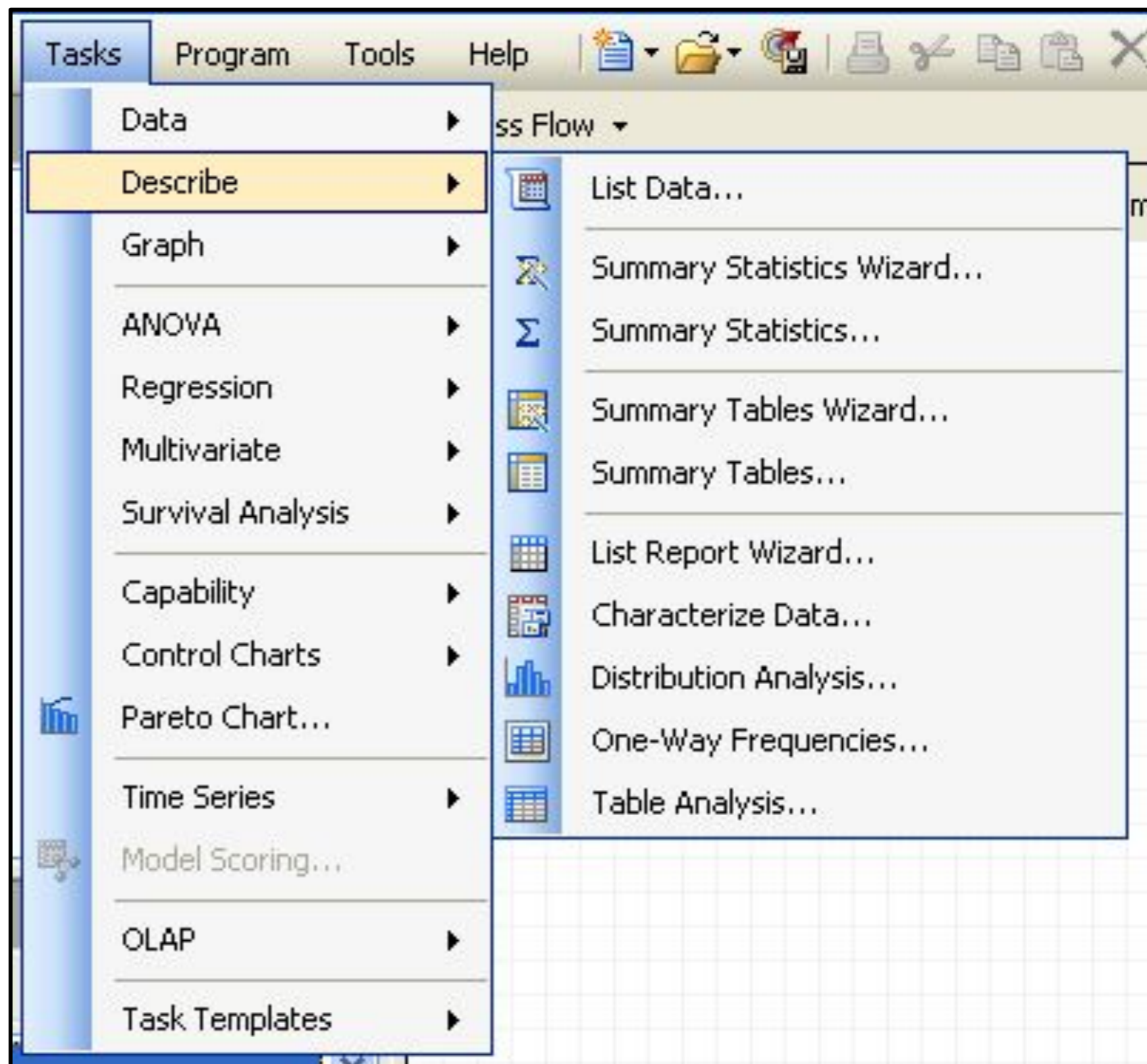
Перечень основных аналитических
Задач группы Data – преобразований с
данными:

- Фильтр и сортировка
- Редактор запроса
- Присоединение таблиц
- Создание формата
- Транспонирование, Split, Stack
- Случайная выборка
- Ранг
- Стандартизация данных
- Свойства атрибутов
- Сравнение данных



Группа узлов Describe (описание данных)

- Списочный отчет
- Отчет сводной статистики
- Сводная таблица
- Характеристика данных
- Анализ распределений
- Односторонний анализ частот
- Кросс - таблицы



Ассоциация задач с данными

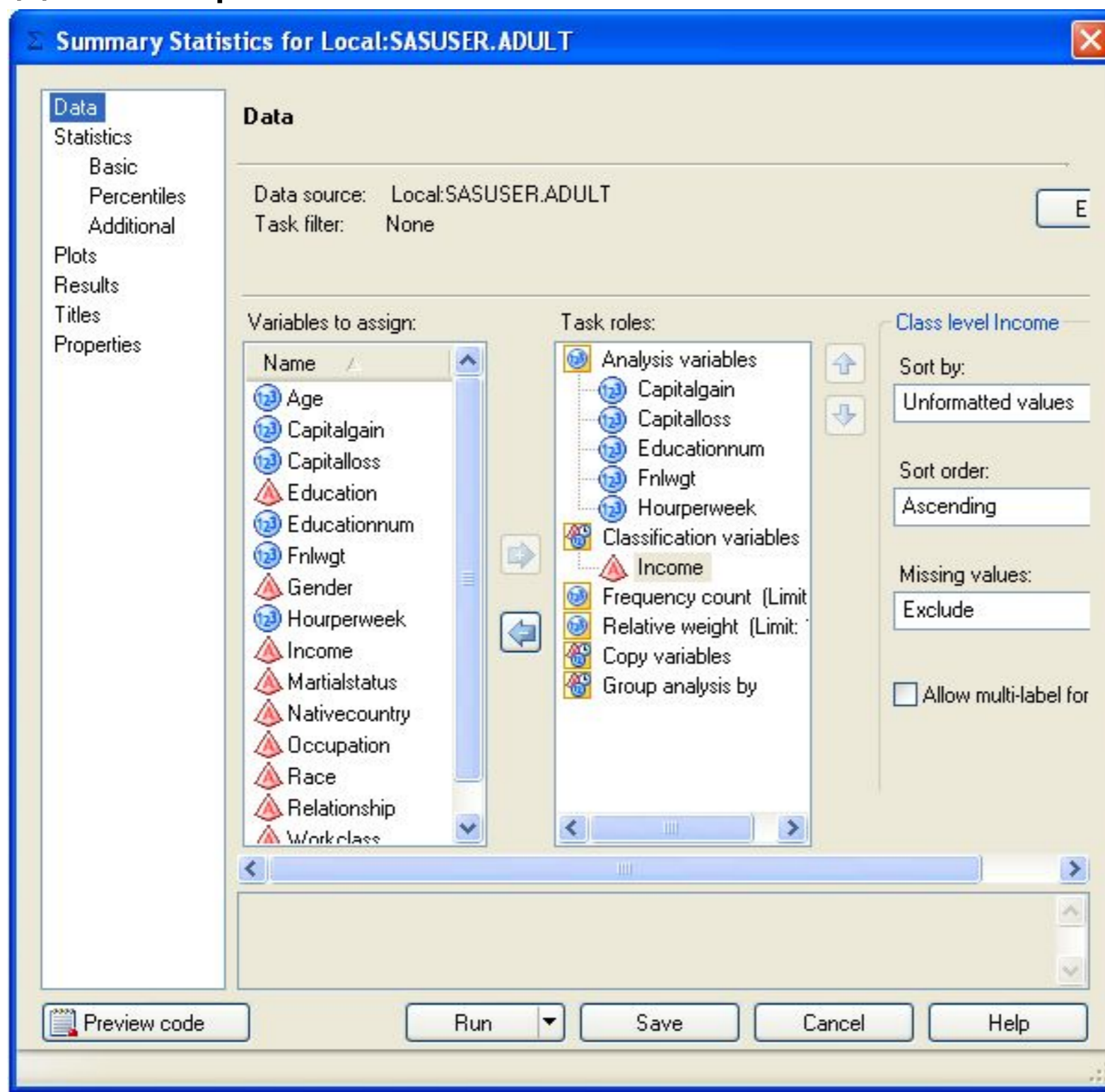
- Задачи также можно выбирать из специального подменю при просмотре данных.

The screenshot shows a data analysis software interface. At the top, there is a toolbar with buttons for 'Filter and Sort', 'Query Builder', 'Data', 'Describe', 'Graph', 'Analyze', 'Export', and 'Send To'. The 'Describe' menu is open, showing options like 'List Data...', 'Summary Statistics Wizard...', 'Summary Statistics...', 'Summary Tables Wizard...', 'Summary Tables...', 'List Report Wizard...', 'Characterize Data...', 'Distribution Analysis...', 'One-Way Frequencies...', and 'Table Analysis...'. The 'One-Way Frequencies...' option is highlighted. Below the toolbar, there is a data table with columns 'Customer_ID', 'Customer_Coun', 'Name', and 'Custor'. The table contains 15 rows of data.

	Customer_ID	Customer_Coun	Name	Custor
1	4	US	James	
2	5	US	Sandrina	
3	9	DE	Cornelia	
4	10	US	Karen	
5	11	DE	Elke	
6	12	US	David	
7	13	DE	Markus	
8	16	DE	Ulrich	
9	17	US	Jimmie	
10	18	US	Tonie	
11	19	DE	Oliver S.	
12	20	US	Michael	
13	23	US	Tulio	
14	24	US	F Robyn Klem	Robyn
15	27	US	F Cynthia McCluney	Cynthia

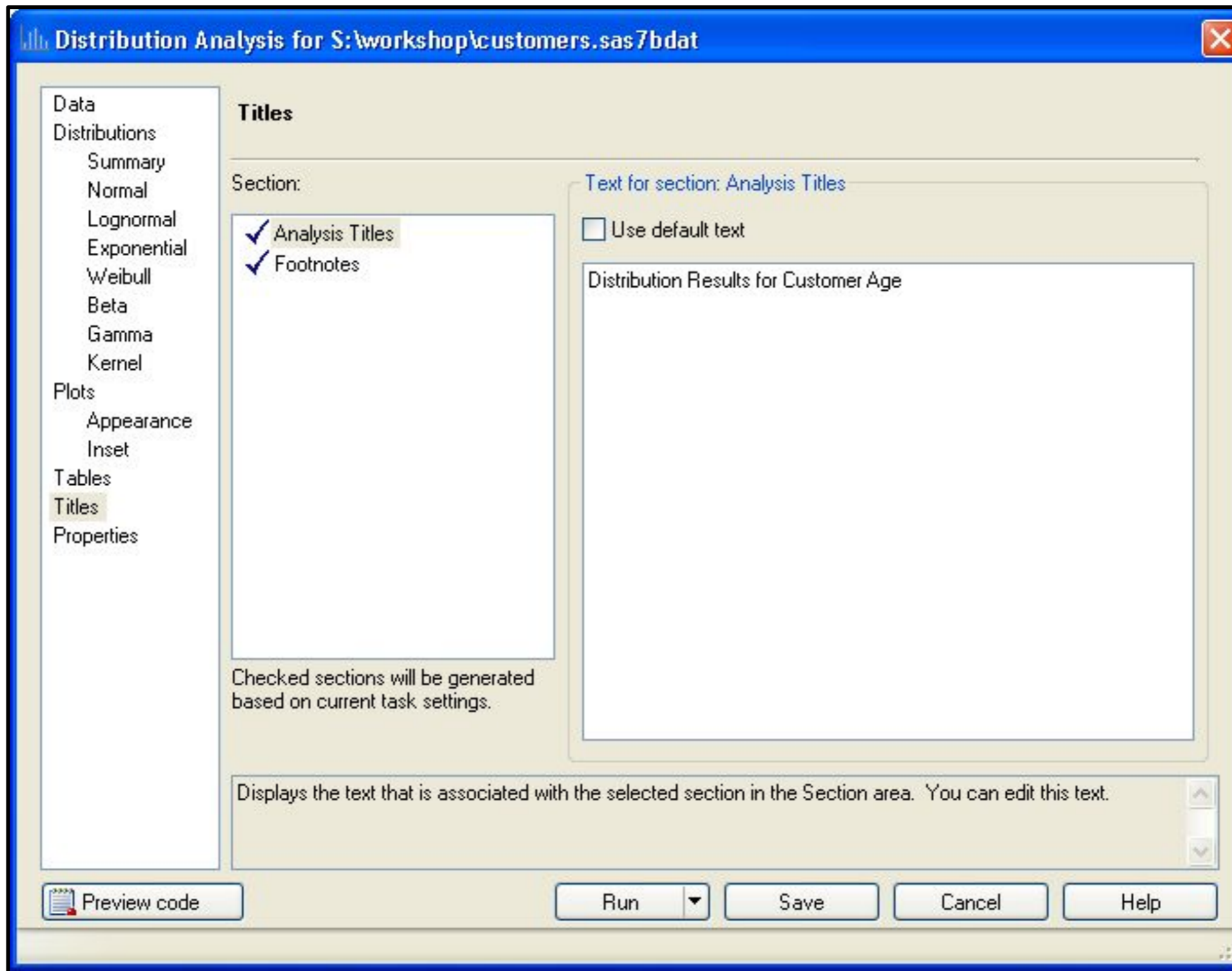
Задачи: Выбор анализируемых переменных

- Каждое диалоговое окно настроек для задач имеет меню настроек входных переменных



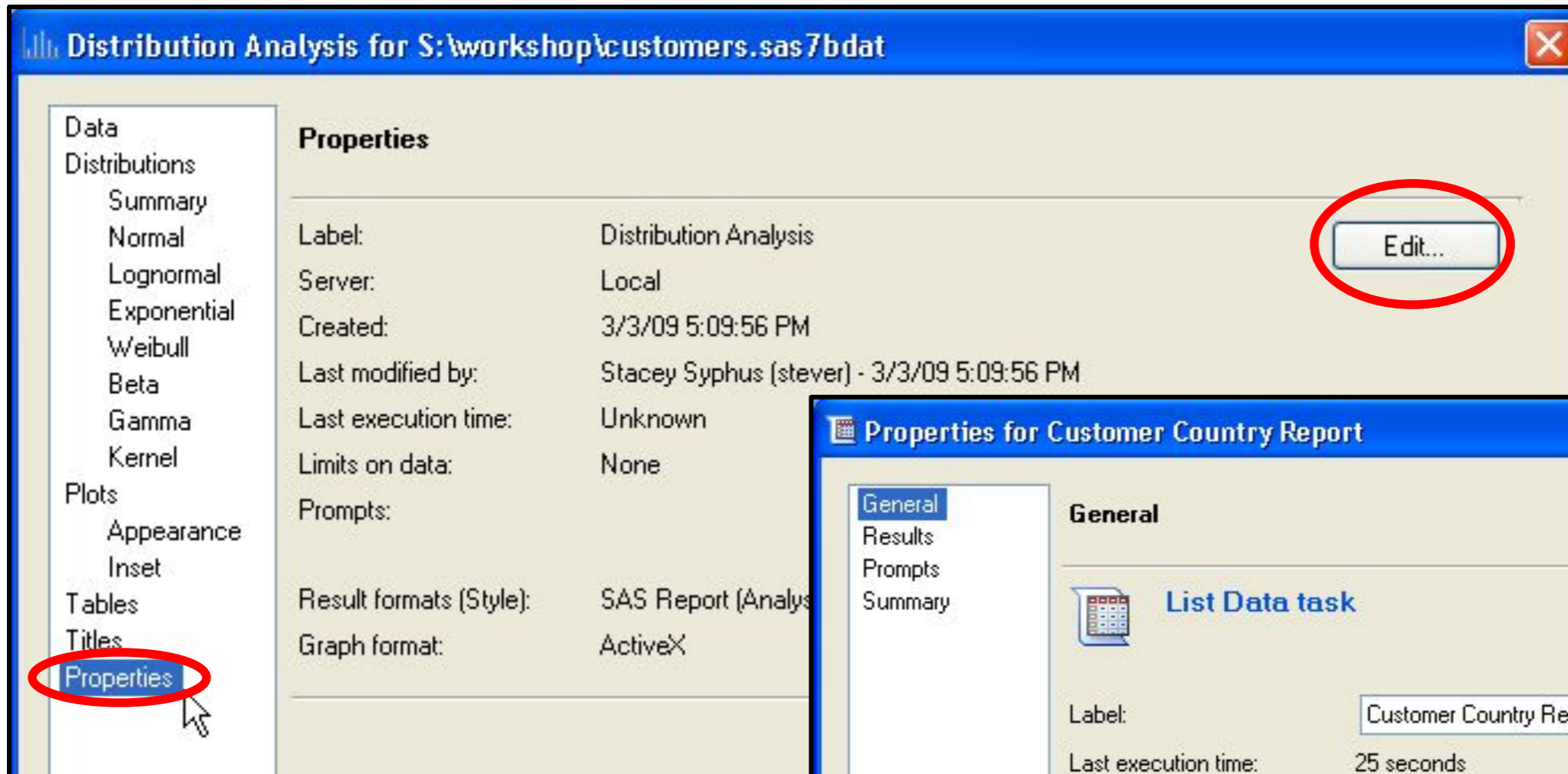
Задачи: Заголовки и подписи

- Выберите вкладку **Titles** чтобы добавить заголовков или подпись к задаче.



Задачи свойства

- В поле свойств можно изменять некоторые доп. настройки вывода для задачи.



Работа с задачами

Код, лог-файлы и выходные данные получают в виде отдельных вкладок, как результат исполнения узла.

The screenshot shows the SAS Enterprise Guide interface. The 'Project Tree' on the left displays a 'Process Flow' containing a 'customers' node and a 'Characterize Data' node. The main window is titled 'Characterize Data' and features a toolbar with the following items: 'Input Data', 'Code', 'Log', 'Output Data (2)', 'Results', 'Refresh', 'Modify Task', 'Export', 'Send To', 'Create', 'Publish', and 'Properties'. The 'Results' tab is active, displaying a 'Summary of Categories Limited to the 30 Most' section. Below this, a table lists variables and their labels:

Variable	Label
Customer_Age_Group	Customer Age Group
Customer_Country	Customer Country

Транспонирование

- За всем тремя SAS Enterprise Guide задачами реструктуризации колонок стоит SAS процедура PROC TRANSPOSE:
 - Split Columns (разбивка колонок)
 - Stack Columns (сшивки колонок)
 - Transpose (транспонирование)

Транспонирование

- Вот как будут выглядеть результаты транспонирования для задачи где требуется вынести, в отдельные колонки, рабочий домашний и сотовый телефон.

Employee_ID	Phone_Type	Country	phone
120106	Home	AU	(02) 2555 5993
120106	Work	AU	(02) 2555 3888
120106	Cell	AU	0461 555 555
120107	Home	AU	(02) 3555 4820
120107	Work	AU	(02) 3555 8333

Employee_ID	Country	Source	Home	Work	Cell
120106	AU	phone	(02) 2555 5993	(02) 2555 3888	0461 555 555
120107	AU	phone	(02) 3555 4820	(02) 3555 8333	
120108	AU	phone	(02) 3555 2315	(02) 3555 6344	

Задача Транспонирование

- Также транспонирование может применять и для сшивки колонок аналогично stack column

	Customer_ID	Month1	Month2	Month3	Month4	Month5	Month6
1	5	213.1	.	478	525.8	394.35	191.79
2	10	188.1	414.09	2876.9	3164.59	2373.4425	169.29
3	11	78.2	70.38
4	12	135.6	.	117.6	129.36	97.02	122.04
5	18	.	.	29.4	32.34	24.255	.
6	24	93	265.8	.	.	.	83.7
7	27	310.7	782.9	.	.	.	279.63
8	31	1484.3	293.3	.	.	.	1335.87
9	34	642.5	.	86.3	94.93	71.1975	578.25

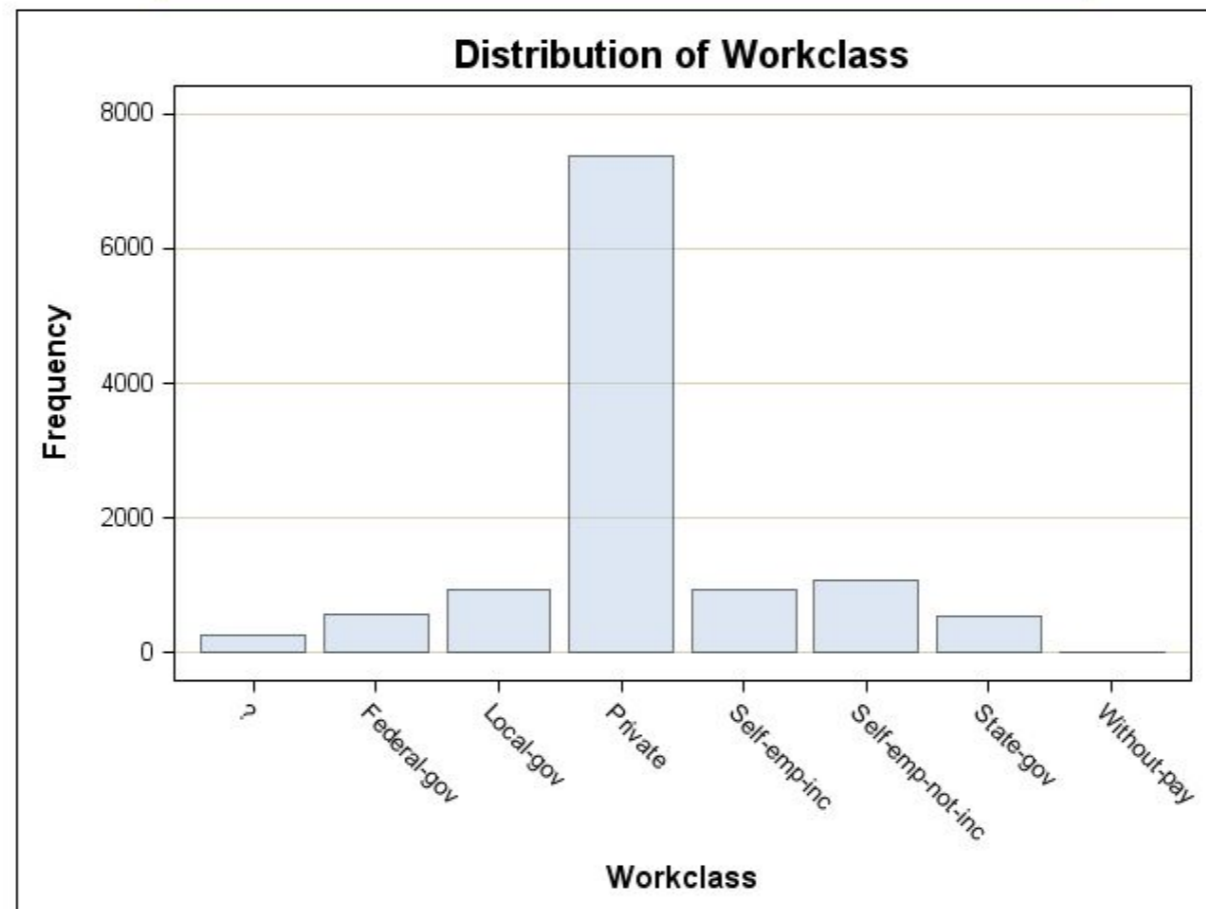
Customer_ID	Month	OrderAmount1
	Month1	213.1
	Month2	.
	Month3	478
	Month4	525.8
	Month5	394.35
	Month6	191.79
10	Month1	188.1
10	Month2	414.09
10	Month3	2876.9
10	Month4	3164.59

Анализ частот

- Односторонний Частотный анализ позволяет создать таблицу и график с указанием кол-ва, процентов, по различным категориям рассматриваемых переменных

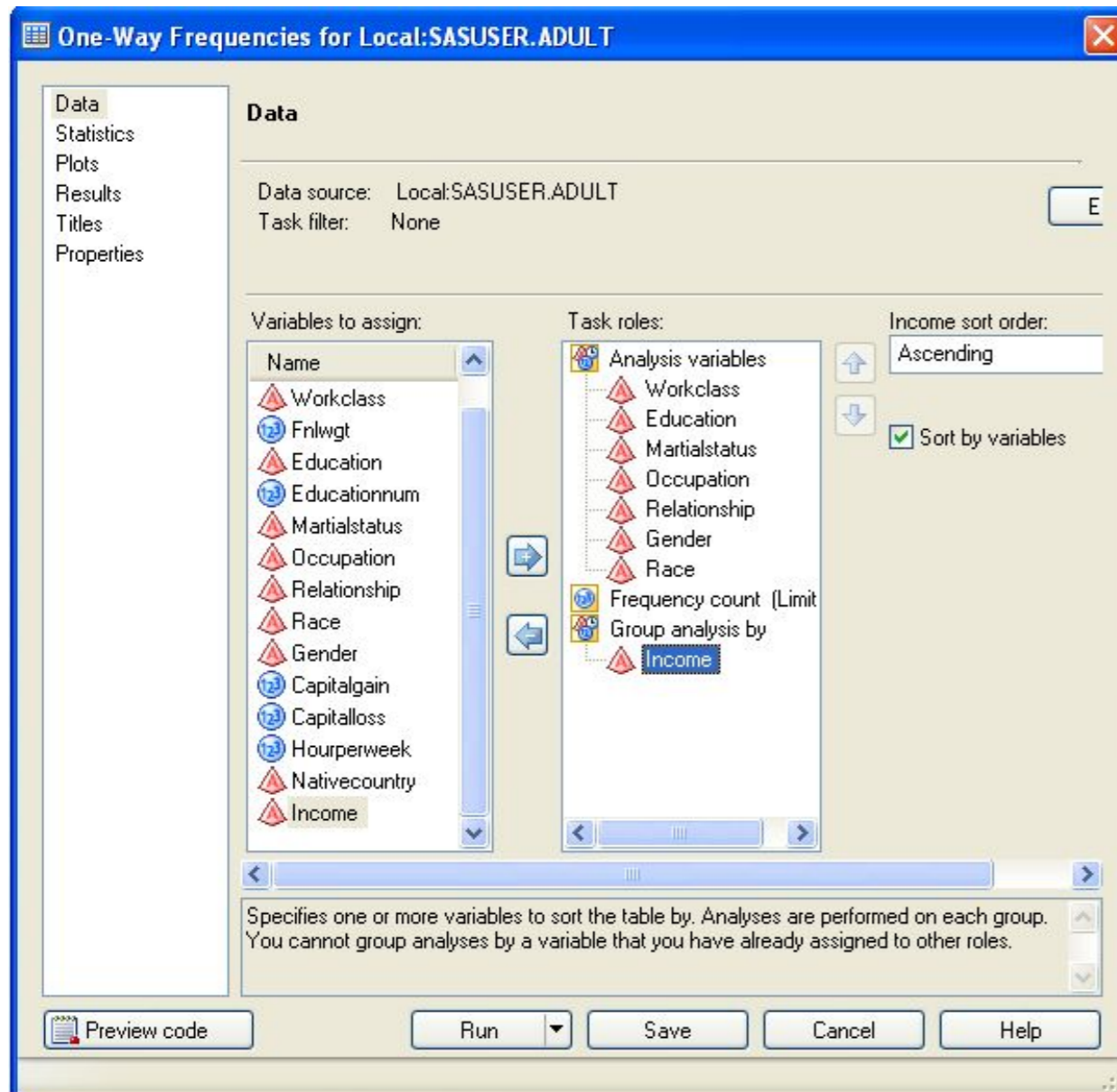
Income=>50K

Workclass	Frequency	Percent	Cumulative Frequency	Cumulative Percent
?	265	2.27	265	2.27
Federal-gov	561	4.80	826	7.07
Local-gov	927	7.93	1753	15.00
Private	7387	63.21	9140	78.21
Self-emp-inc	938	8.03	10078	86.23
Self-emp-not-inc	1077	9.22	11155	95.45
State-gov	530	4.53	11685	99.98
Without-pay	2	0.02	11687	100.00



Роли переменных для Анализа частот.

- В закладке настроек data необходимо указать анализируемые переменные



Income<=50K

One-Way Frequencies Results The FREQ Procedure

Workclass	Frequency	Percent	Cumulative Frequency	Cumulative Percent
?	2534	6.82	2534	6.82
Federal-gov	871	2.34	3405	9.16
Local-gov	2209	5.95	5614	15.11
Never-worked	10	0.03	5624	15.14
Private	26519	71.37	32143	86.51
Self-emp-inc	757	2.04	32900	88.55
Self-emp-not-inc	2785	7.50	35685	96.04
State-gov	1451	3.91	37136	99.95
Without-pay	19	0.05	37155	100.00

Частотный анализ (дополнительные настройки)

- Используя расширенные настройки можно создавать различные ВИДЫ ОТЧЕТОВ.

The image displays two overlapping screenshots of the SAS software interface, illustrating the configuration options for a frequency analysis.

Top Screenshot (Statistics Panel):

- Left Navigation:** Data, **Statistics**, Plots, Results, Titles, Properties.
- Statistics Panel:**
 - Frequency table options:** Includes radio buttons for:
 - Frequencies and percentages with cumulatives
 - Frequencies and cumulative frequencies
 - Frequencies and percentages
 - Frequencies only
 - Binomial proportions:** Includes checkboxes for:
 - Asymptotic test
 - Exact p-values
 - Test proportion:** Input field with value 0.5
 - Confidence level:** Dropdown menu with value 95%

Bottom Screenshot (Results Panel):

- Left Navigation:** Data, Statistics, Plots, **Results**, Titles, Properties.
- Results Panel:**
 - Create data set with frequencies and percentages
 - Add cumulative frequencies and cumulative percentages to data set
 - Analysis Variables:** Input field with value Product_Category
 - Data file for Product_Category:** Input field with value Local:SASUSER.OneWayFreqOfProduct_CategoryInPR

Виды форматов для выходных результатов

- По умолчанию, SAS Enterprise Guide строить отчеты в формате SAS Report. Однако вы также можете создавать HTML, PDF, RTF, или текстовые отчеты.

The diagram illustrates four different output formats for the same data set, titled "Number of Products per Category".

SAS Report Format: A table with a header "Product Category" and columns "Product_Category", "Frequency", and "Percent".

Product_Category	Frequency	Percent
Assorted Sports Articles	68	14.14
Children Sports	27	5.61
Clothes	108	22.45
Golf	56	11.64

PDF Format: A table with a header "Product Category" and columns "Product_Category", "Frequency", and "Percent".

Product_Category	Frequency	Percent
Assorted Sports Articles	68	14.14
Children Sports	27	5.61
Clothes	108	22.45
Golf	56	11.64

Word Document Format: A table with a header "Product Category" and columns "Product_Category", "Frequency", and "Percent".

Product_Category	Frequency	Percent
Assorted Sports Articles	68	14.14
Children Sports	27	5.61
Clothes	108	22.45
Golf	56	11.64
Indoor Sports	15	3.12

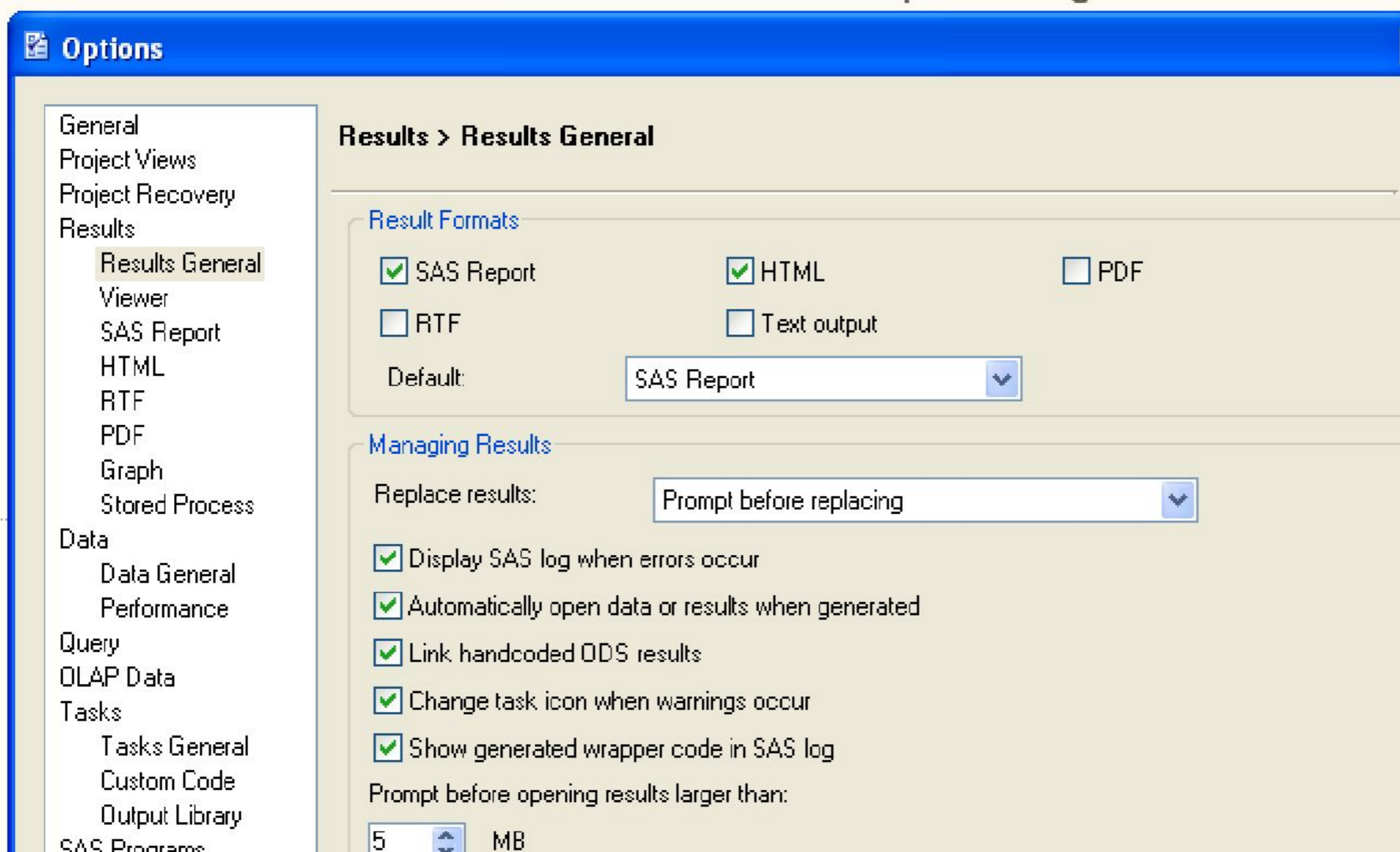
Text File Format: A plain text table with a header "Product Category" and columns "Product_Category", "Frequency", and "Percent".

Product_Category	Frequency	Percent
Assorted Sports Articles	68	14.14
Children Sports	27	5.61
Clothes	108	22.45
Golf	56	11.64
Indoor Sports	15	3.12
Outdoors	87	18.09
Racket Sports	14	2.91
Running - Jogging	25	
Shoes	34	
Swim Sports	7	
Team Sports	23	
Winter Sports	17	

Icons for each format: Internet Explorer (SAS Report), Adobe PDF (PDF), Microsoft Word (Word Document), and a document icon (Text File).

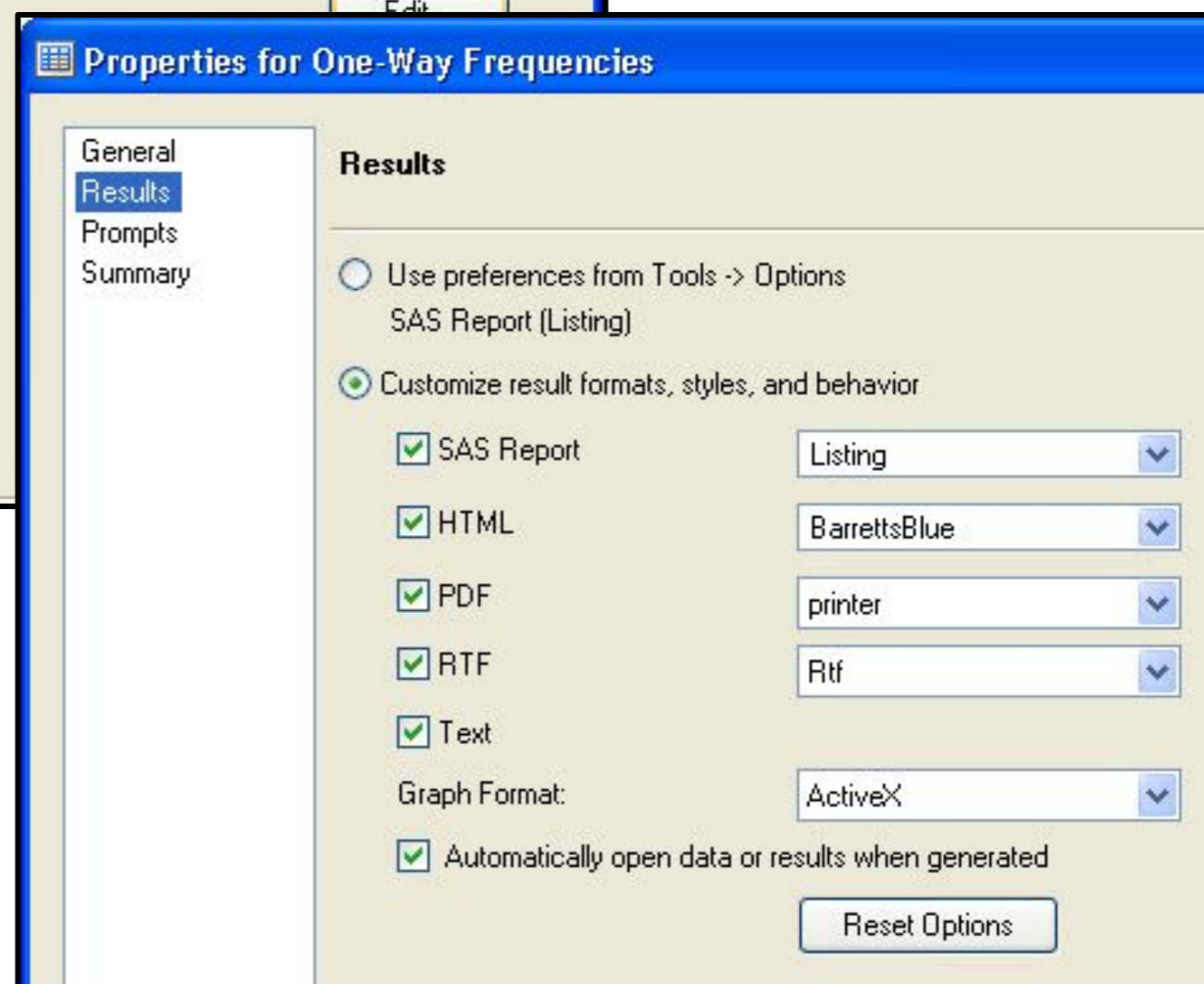
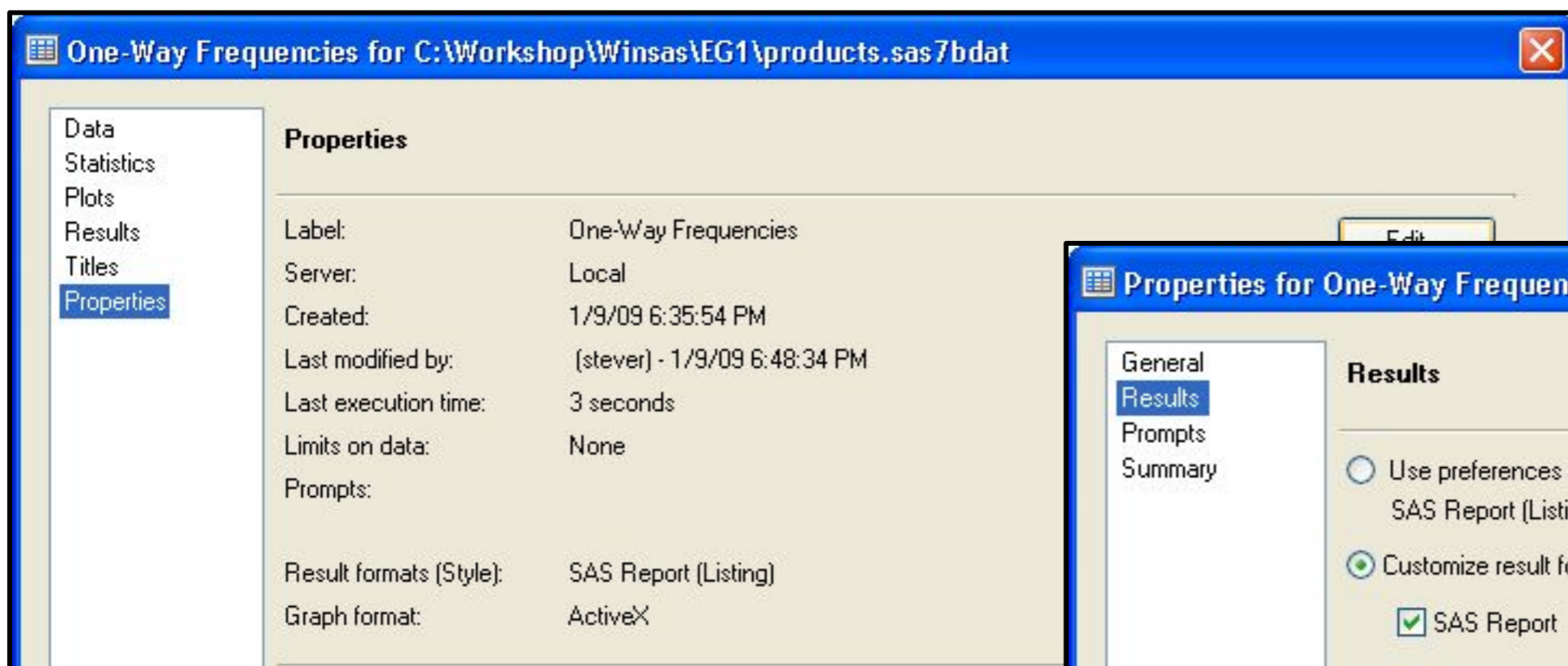
Настройка выходных форматов (глобальная)

- Чтобы поменять типы отчетов по умолчанию откройте **Tools** → **Options** → **Results General**. И выберите требуемый вид создаваемых отчетов.



Локальная настройка формата отчета.

- В свойствах конкретного узла, также можно поменять формат выходного отчета, однако это будет локальная установка формата

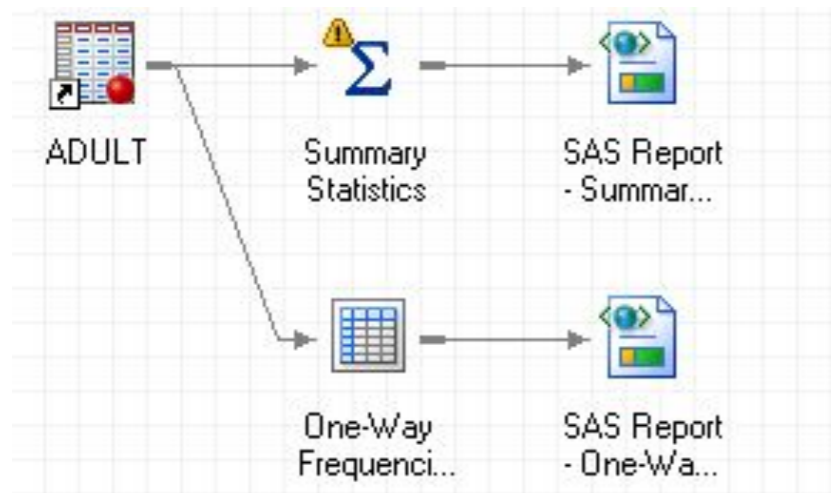
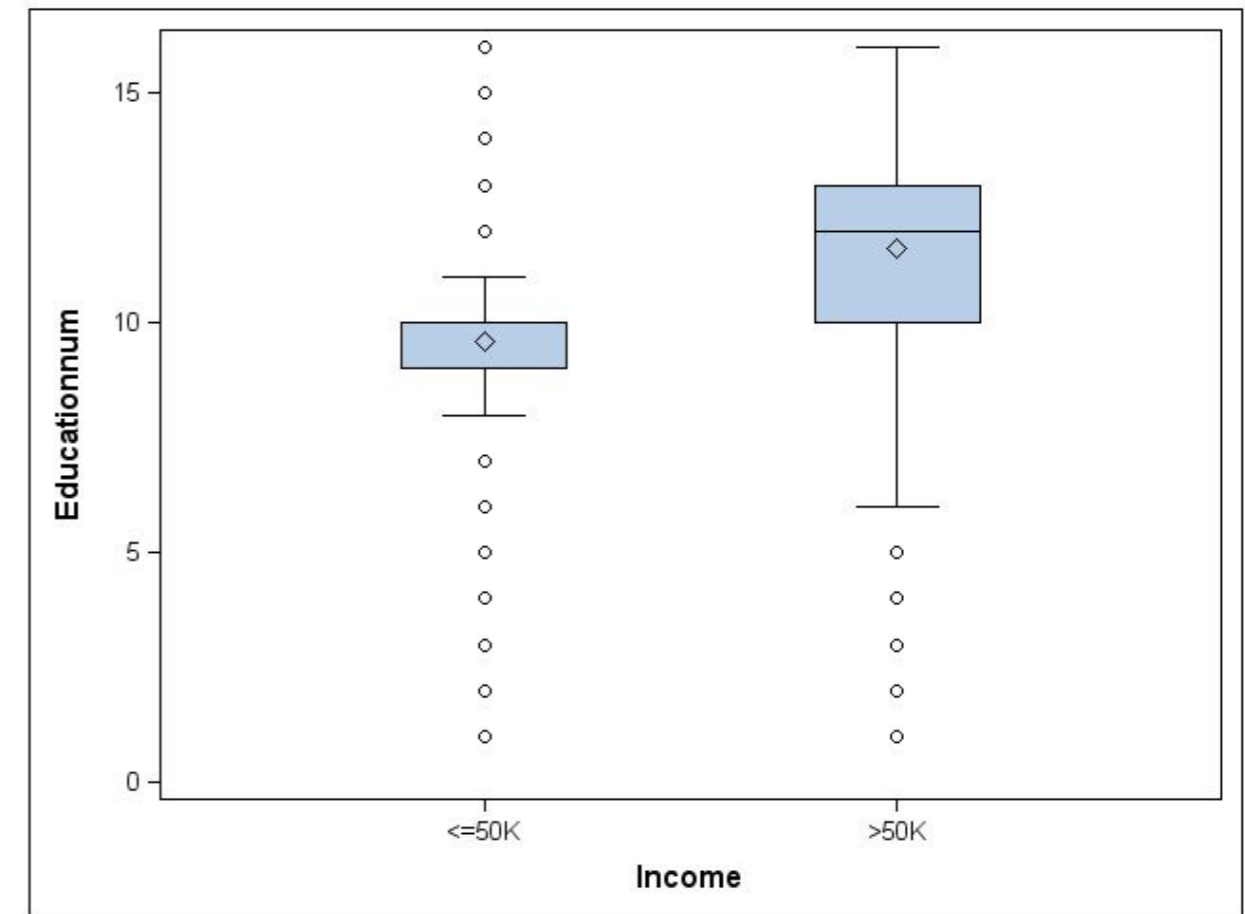


Задача сводной статистики

- Сводная статистика позволяет вычислять базовые наборы статистических показателей, в том числе и по группам. Результаты можно сохранять в виде графических отчетов, и таблиц с данными

Summary Statistics
Results
The MEANS Procedure

Income	N Obs	Variable	Mean	Std Dev	Minimum	Maximum	N
<=50K	37155	Capitalgain	147.0103082	936.7536781	0	41310.00	37155
		Capitalloss	54.1519311	313.3200048	0	4356.00	37155
		Educationnum	9.5984928	2.4376728	1.0000000	16.0000000	37155
		Fnlwgt	190039.57	106577.60	12285.00	1490400.00	37155
		Hourperweek	38.8400484	12.3568494	1.0000000	99.0000000	37155
>50K	11687	Capitalgain	4042.24	14756.77	0	99999.00	11687
		Capitalloss	193.5289638	593.2116118	0	3683.00	11687
		Educationnum	11.6028065	2.3826241	1.0000000	16.0000000	11687
		Fnlwgt	188470.57	102442.73	13769.00	1226583.00	11687
		Hourperweek	45.4528964	11.0911762	1.0000000	99.0000000	11687



Настройка сводной статистики

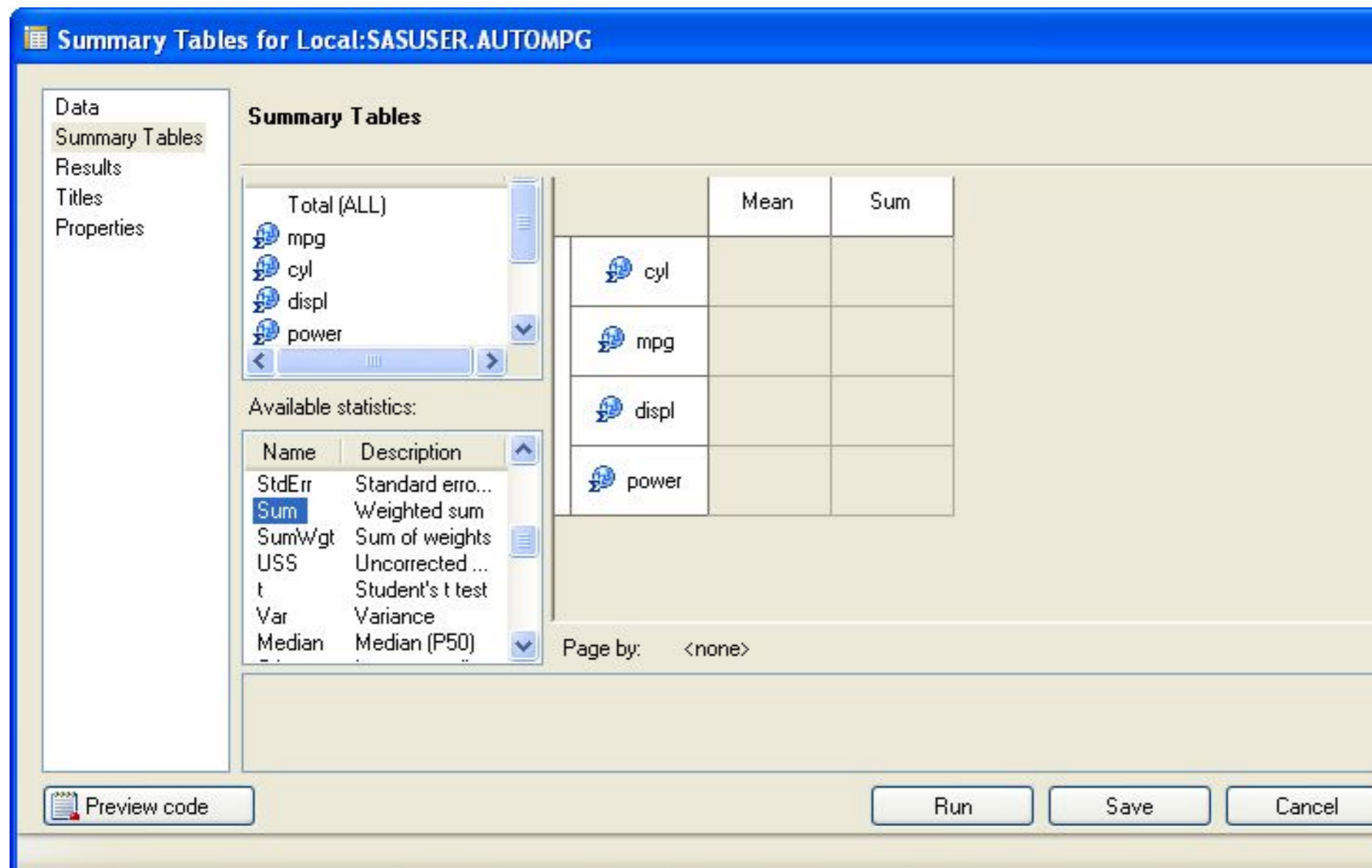
- Сводная статистика

The image displays three overlapping screenshots from the SAS software interface:

- Top-left:** A menu view showing the 'Summary Statistics Wizard...' option highlighted with a red rectangle under the 'Describe' menu.
- Top-right:** The 'Summary Statistics2' dialog box for the dataset 'S:\workshop\orion_profit.sas7bdat'. The 'Data' tab is active, showing a list of variables on the left and 'Customer_Age' selected as the 'Analysis variable' on the right.
- Bottom:** The 'Assign variables to roles' step of the wizard. It shows 'Customer_Age' assigned to 'Summary statistics of (Analysis variable)', 'Customer_Country' assigned to 'For each value of (Classification variable)', and an empty field for 'Separate tables for values of (Group variable)'. The 'Advanced...' button is visible at the bottom right.

Сводные таблицы

- Сводные таблицы, позволяют создавать табличные отчеты специального вида (Pivot tables)



Сводная таблица, расширенные настройки

- Используя сводные таблицы в расширенном режиме, можно модифицировать метки, шрифты и форматы в таблице.

The screenshot shows the 'Summary Tables' dialog box in SAS. The window title is 'Summary Tables2 for S:\workshop\orion_profit.sas7bdat'. On the left, a navigation pane shows 'Data', 'Summary Tables', 'Results', 'Titles', and 'Properties'. The main area is divided into 'Available variables:' and 'Available statistics:'. Under 'Available variables:', 'Total (ALL)', 'Profit', 'Order_Type', and 'Customer_Country' are listed. Under 'Available statistics:', a list of statistical functions like CSS, CV, Max, Mean, Min, N, NMiss, PctN, and PctSum is shown. The 'Preview:' area displays a table structure with columns for 'Order_Type', 'Profit', and 'Customer_Country', and rows for 'Total' and 'Sum'. A context menu is open over the table, with options: 'Remove Cells', 'Table Properties...', 'Box Area Properties...', 'Heading Properties...' (highlighted), 'Class Level Properties...', and 'Data Value Properties...'. At the bottom, there are buttons for 'Run', 'Save', 'Cancel', and 'Help', along with a 'Preview code' button.

Сводная таблица, модификация отображения

The image displays three overlapping 'Table Properties' dialog boxes from a software application. The top-left dialog is on the 'Format' tab, showing categories like 'Numeric' and 'Date/Time', and formats such as 'BESTw.d' and 'COMMAw.d'. It also includes input fields for 'Overall width' (8) and 'Decimal places' (2). The top-right dialog is on the 'Font' tab, showing font settings for 'Arial' in 'Bold' style with a size of '4'. It includes a 'Preview' section with a dark blue background and the text 'Sample Text'. The bottom dialog is on the 'General' tab, showing the 'Label for missing values' set to 'Unknown' and three checkboxes: 'Include class levels with no data.' (unchecked), 'Include missing values as valid levels for class variables.' (checked), and 'Remove all class headings from the row dimension.' (unchecked).

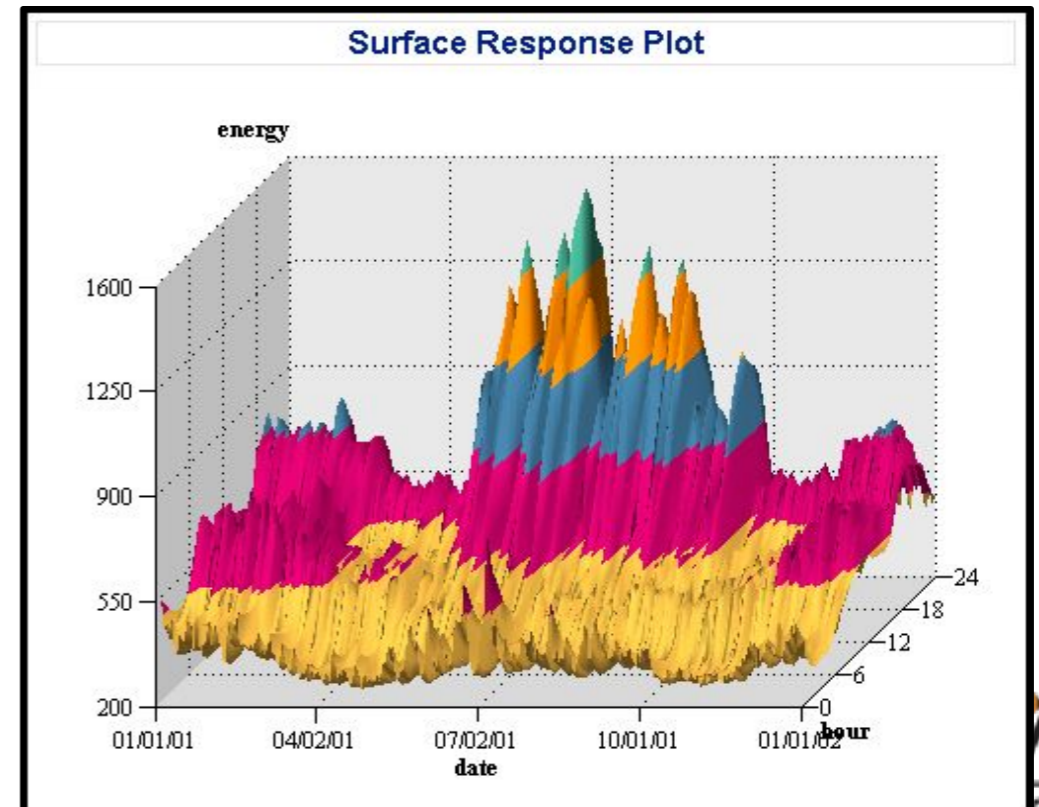
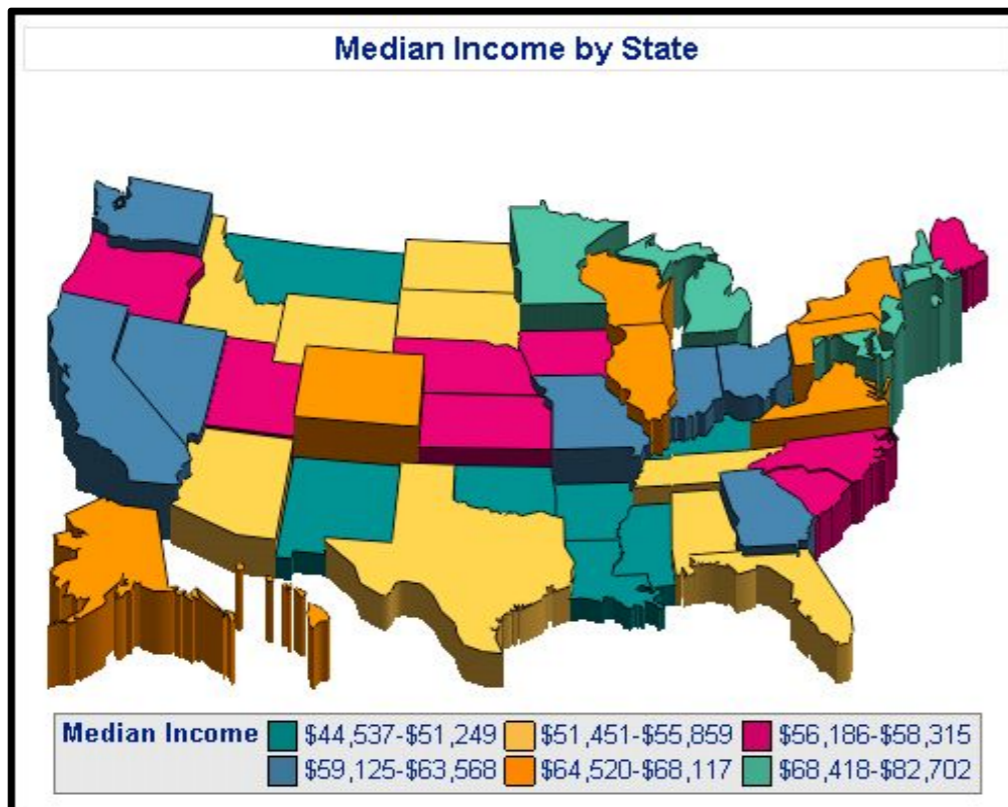
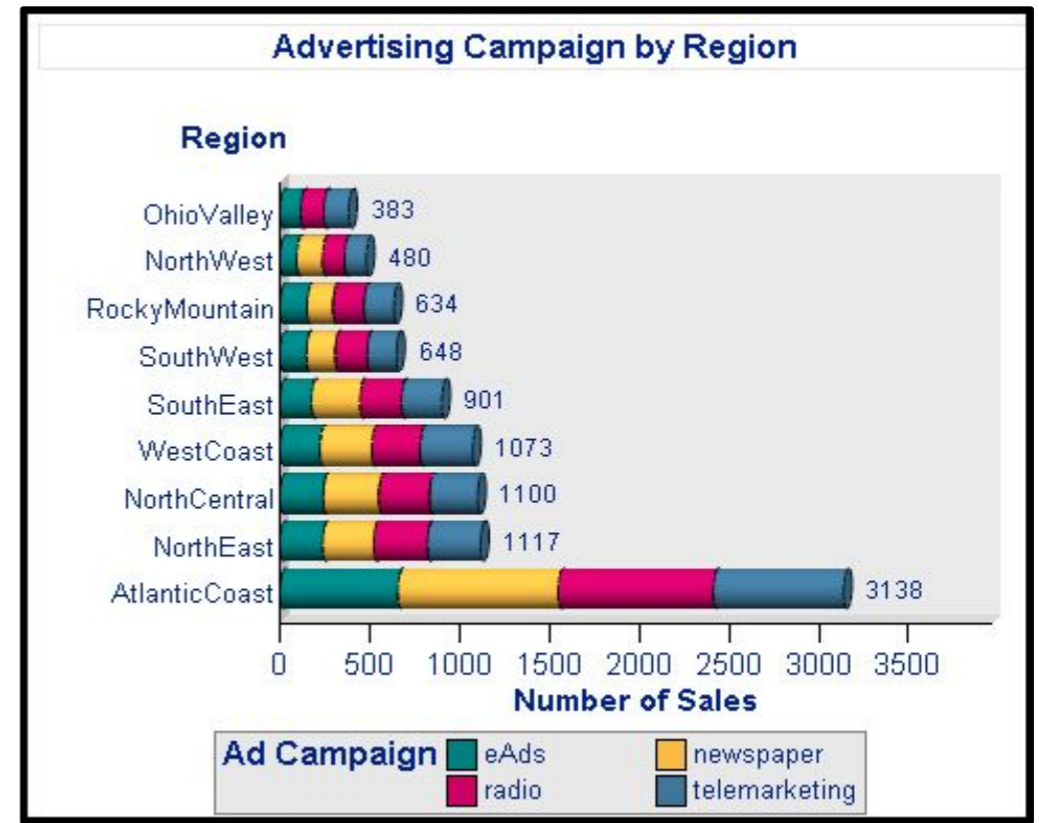
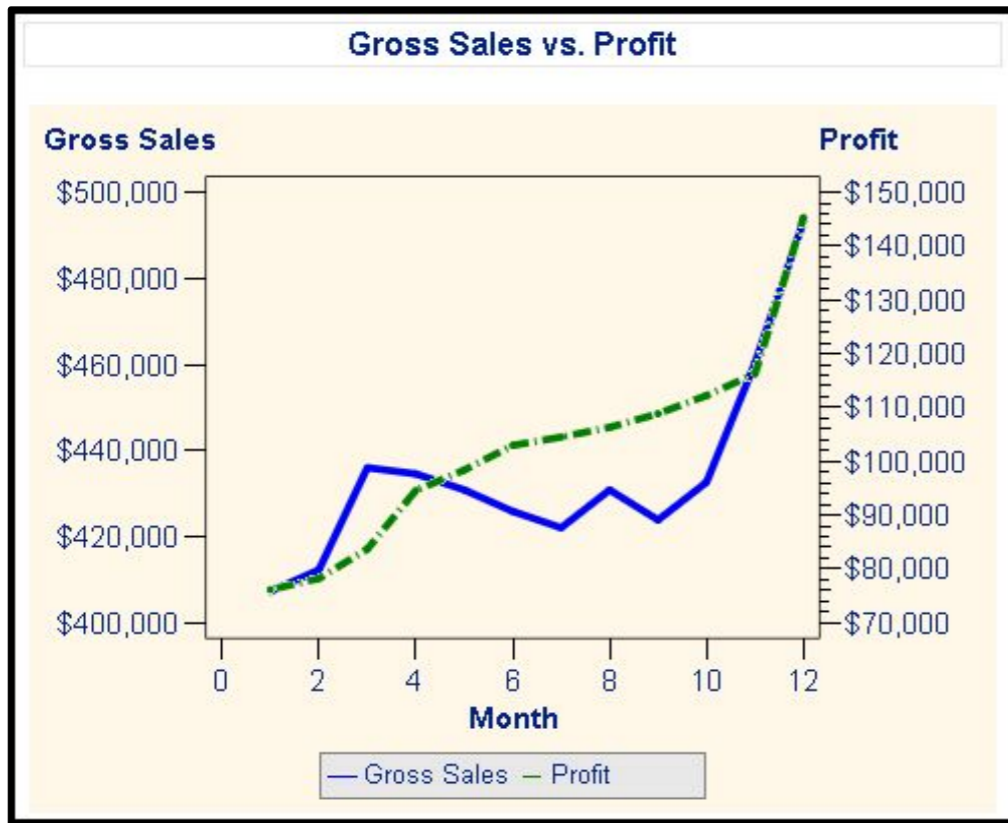
Раздел 5: Создание графических отчетов

5.1. Типы графических отчетов в SAS Enterprise Guide.

5.2. Форматы для графических отчетов SAS Enterprise Guide.

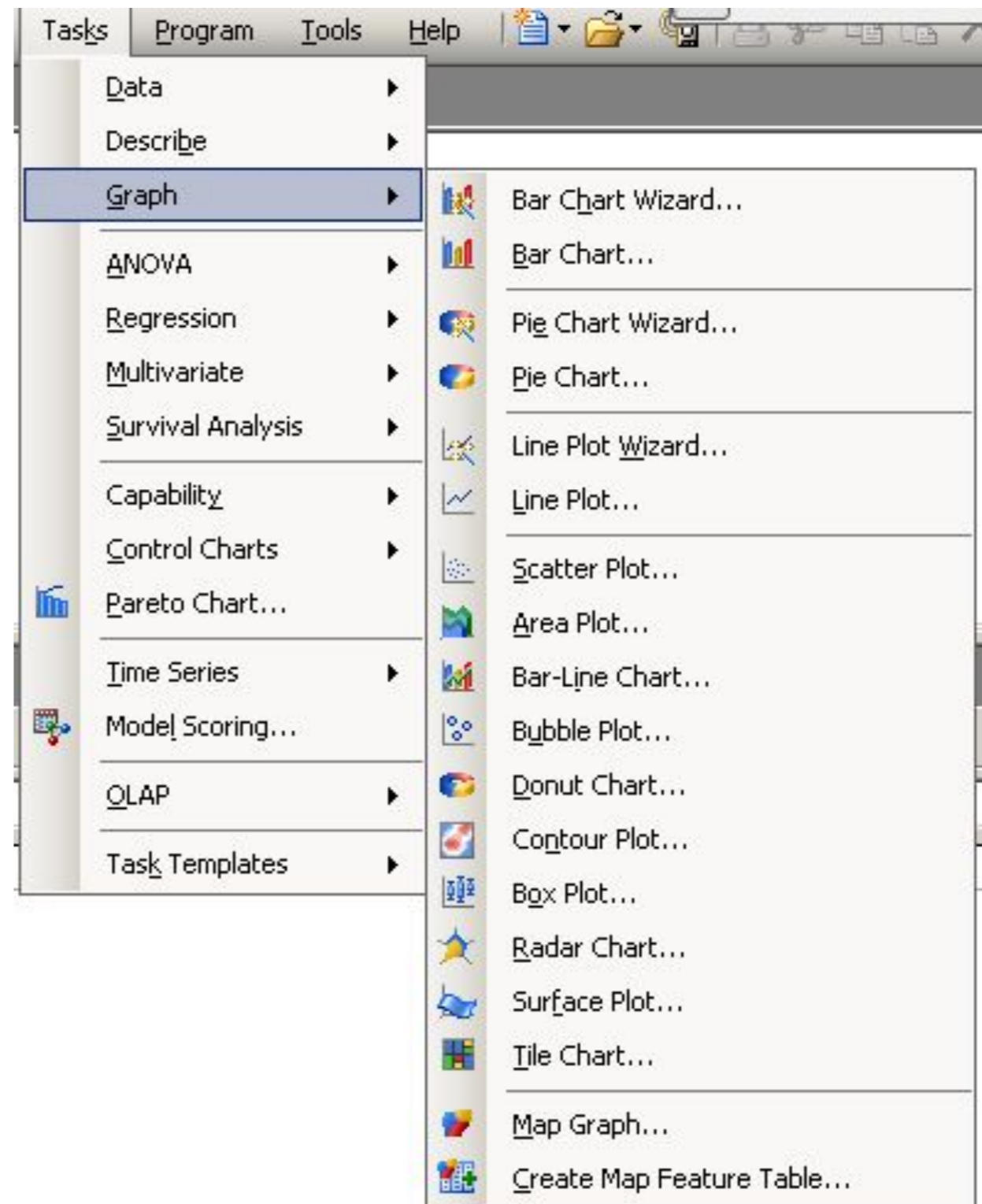
5.3. Создание гистограммы и прочих отчетов

Графические отчеты в SAS Enterprise Guide



Задачи группы Graph – графические отчеты

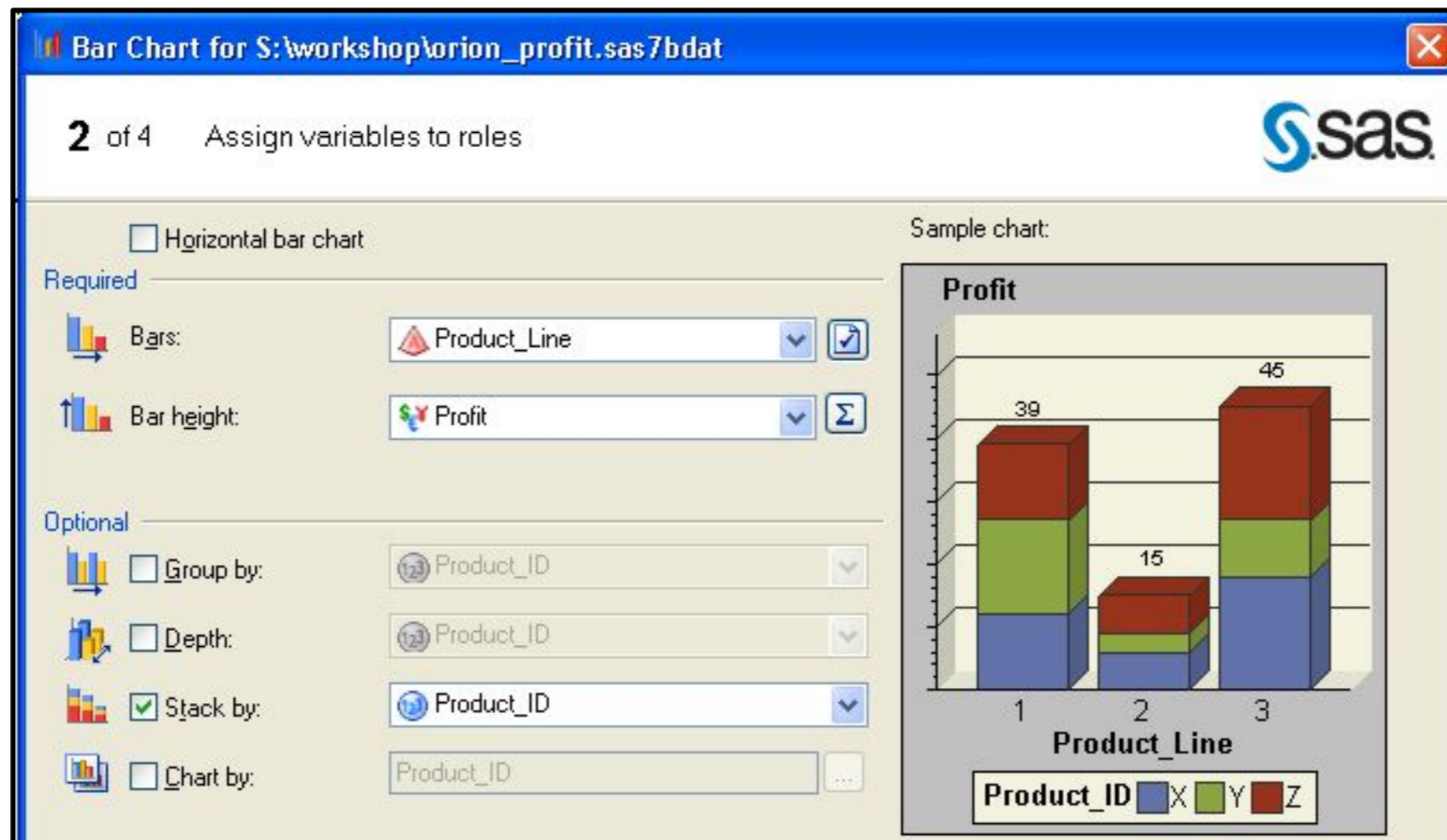
- Гистограмма
- Круговая диаграмма
- Линейный график
- Точечный график
- График площадей
- Пузырьковый график
- Секторный график
- Тепловая карта
- Ящичковая диаграмма
- Радарный график
- Поверхность
- График, объема (Tile)
- Карты



Гистограмма

Настройки гистограмм позволяют:

- 1.Выбирать и фильтровать данные
- 2.Назначать анализируемые переменные
- 3.Управлять внешним видом отчета
- 4.Задавать подписи и заголовки



Гистограмма, настройка в задаче

Bar Chart for Local:EGDATA.TOURS

Bar Chart
Data
Appearance
Bars
Layout
Axes
General
Horizontal Axis
Axis
Vertical Axis
Axis
Major Ticks
Minor Ticks
Reference Lines
Chart Area
Advanced
Titles
Properties

Data

Data source: Local:EGDATA.TOURS
Task filter: None

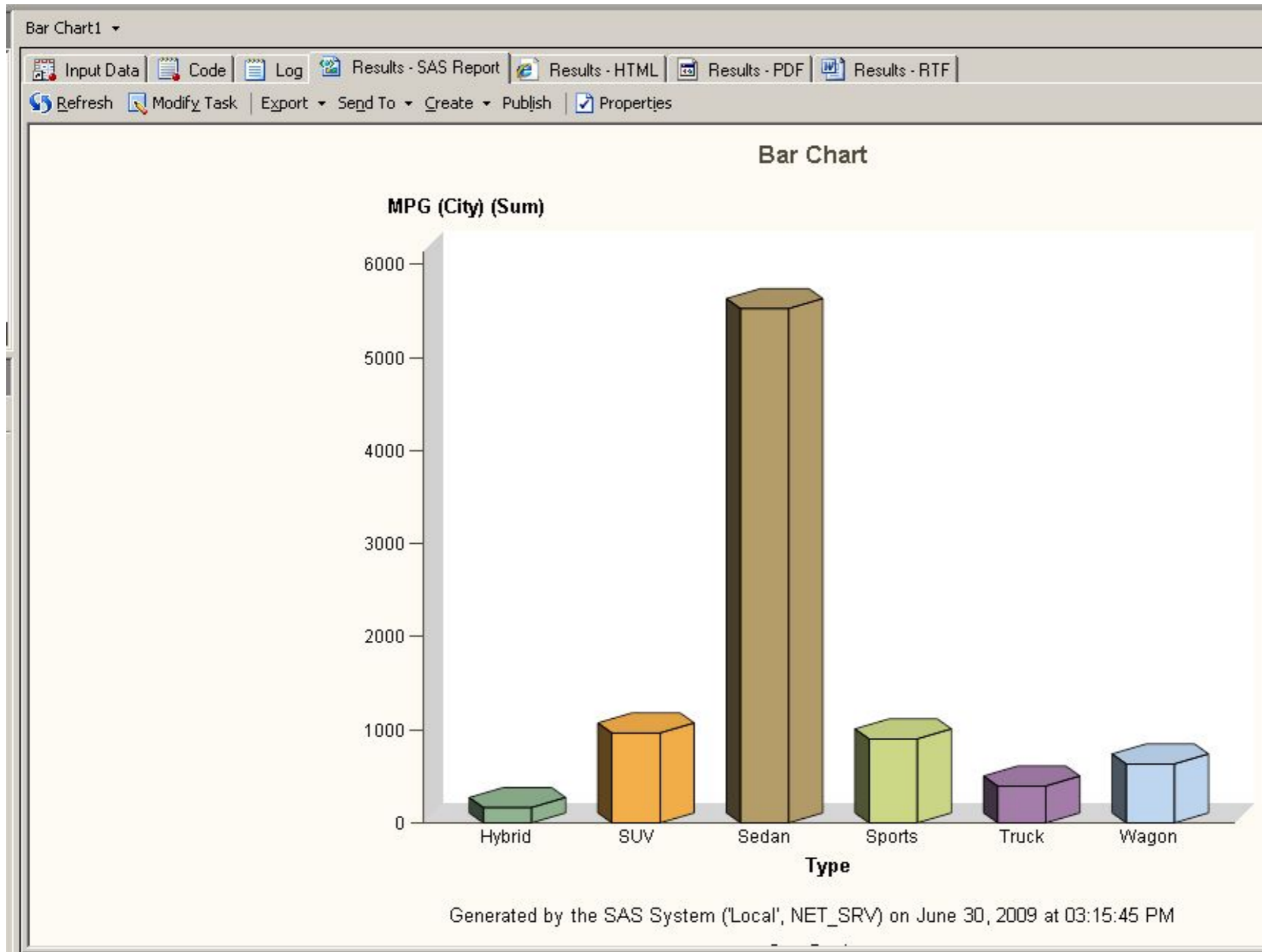
Columns to assign:

Name
Volcano
Departs
Days
Price
Difficulty

Task roles:

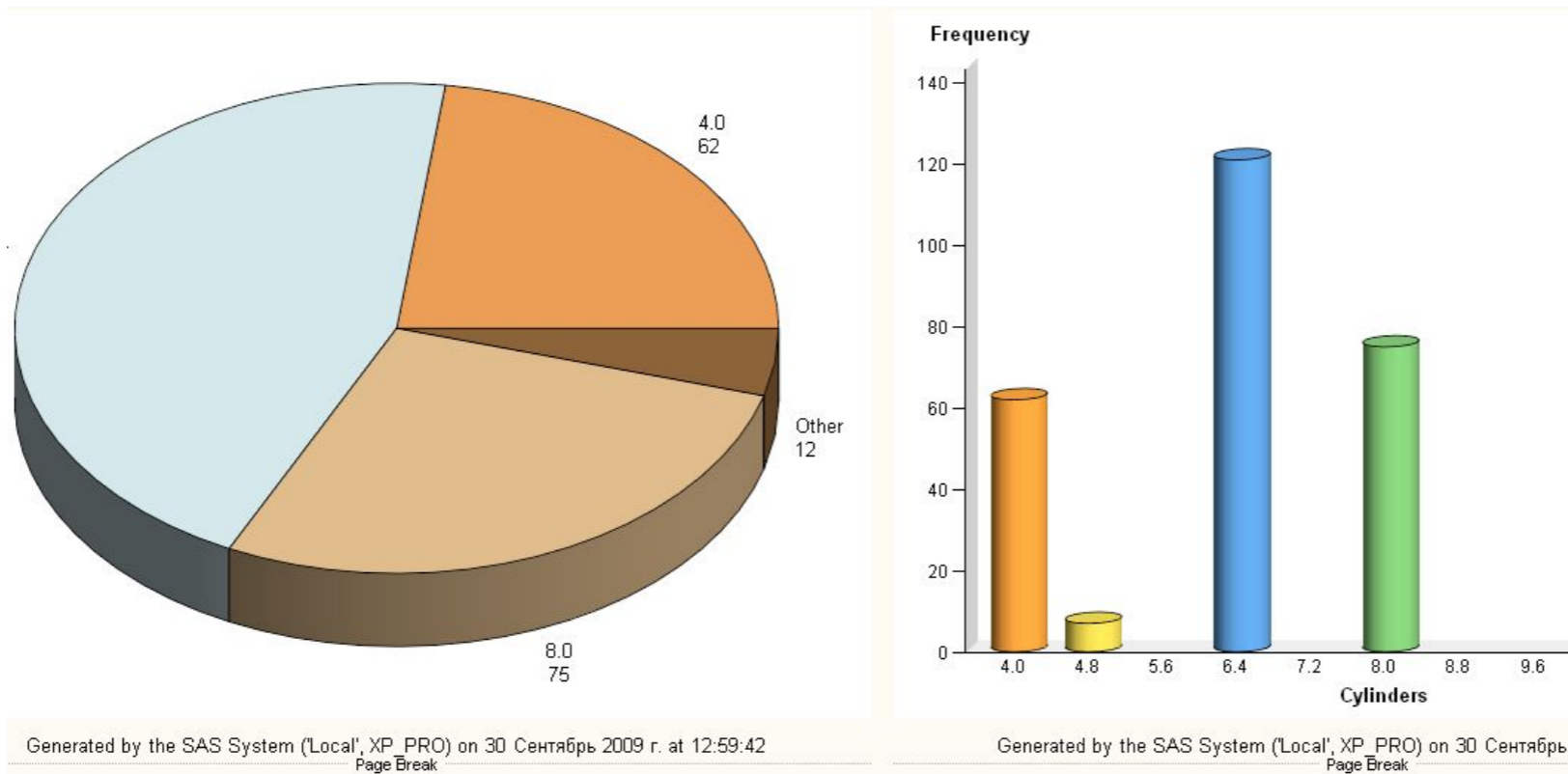
- Column to chart (Limit: 1)
 - Volcano
- Sum of (Limit: 1)
 - Price
- Group charts by

Гистограмма. Результаты



Формат SAS Report

- Формат SAS Report позволяет комбинировать различные результаты анализа в один сводный отчет. Этот отчет можно просмотреть во всех SAS приложениях или экспортировать в HTML, XML, или PDF файл.



Summary Statistics

Results

Origin	N Obs	Variable	Label	Mean	Std Dev	Minimum	Maximum	N
Europe	123	EngineSize	Engine Size (L)	3.2065041	1.0371572	1.6000000	6.0000000	123
		Weight	Weight (LBS)	3680.72	579.5768708	2524.00	5423.00	123
		Cylinders		6.2357724	1.6795847	4.0000000	12.0000000	123
		Horsepower		251.8943089	80.7388842	100.0000000	493.0000000	123
USA	147	EngineSize	Engine Size (L)	3.6428571	1.1942212	1.6000000	8.3000000	147
		Weight	Weight (LBS)	3769.95	855.3055237	2348.00	7190.00	147
		Cylinders		6.1088435	1.5310899	4.0000000	10.0000000	147
		Horsepower		212.8231293	63.7486181	103.0000000	500.0000000	147

Раздел 6: Плагин для Microsoft Office

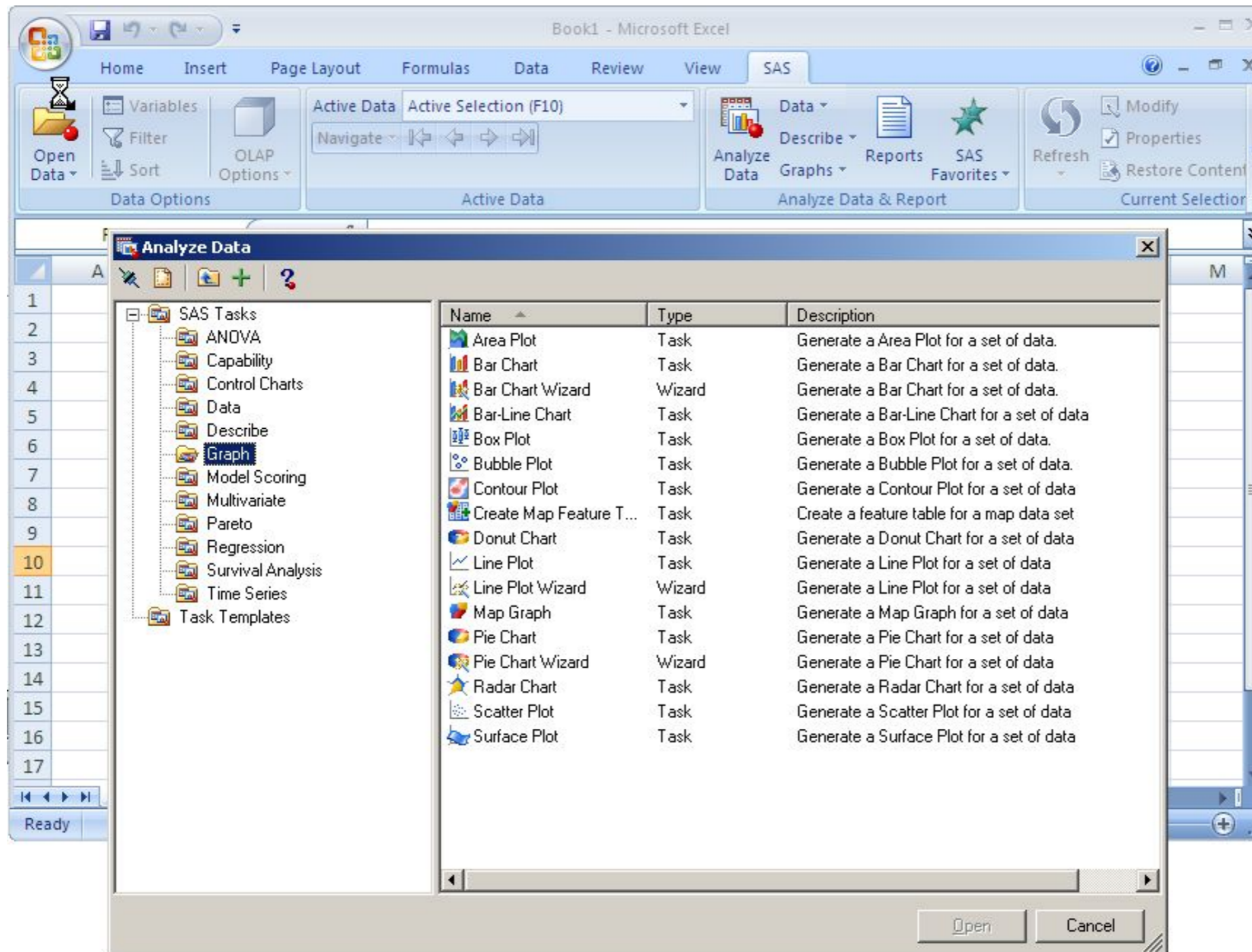
6.1. Возможности плагина

SAS Плагин для Офиса.

- Плагин позволяет
 - Получить доступ к данным SAS и других источников через SAS
 - Обработать данные с помощью инструментов SAS (аналогично EG)
 - Изменить данные удаленных источников

SAS Плагин для Офиса.

- Приложение для работы с SAS в оффисе, позволяет запускать аналитические задачи SAS, непосредственно в MS Excel, Word, Power Point



Импорт данных в Excel

- С помощью SAS плагина, можно загружать в Excel, SAS таблицы, и затем аналитическими задачами обрабатывать их.

The screenshot shows the Microsoft Excel interface with the SAS plugin. The 'SAS' ribbon is active, displaying various data analysis tools. A 'Modify Data Source' dialog box is open, showing the 'Local:SASHELP.CARS' dataset. The 'Available' list includes variables like 'Invoice', 'Engine Size (L)', 'Cylinders', 'Horsepower', 'MPG (City)', 'MPG (Highway)', 'Weight (LBS)', 'Wheelbase (IN)', and 'Length (IN)'. The 'Selected' list includes 'Make', 'Model', 'Type', 'Origin', and 'DriveTrain'. The 'Output Location' section is set to 'New Worksheet' with the name 'SASHELP.CARS'. The 'Update' button is visible, and the status bar shows 'Results: 428 of 428 rows'.

Результаты анализа в SAS плагине.

