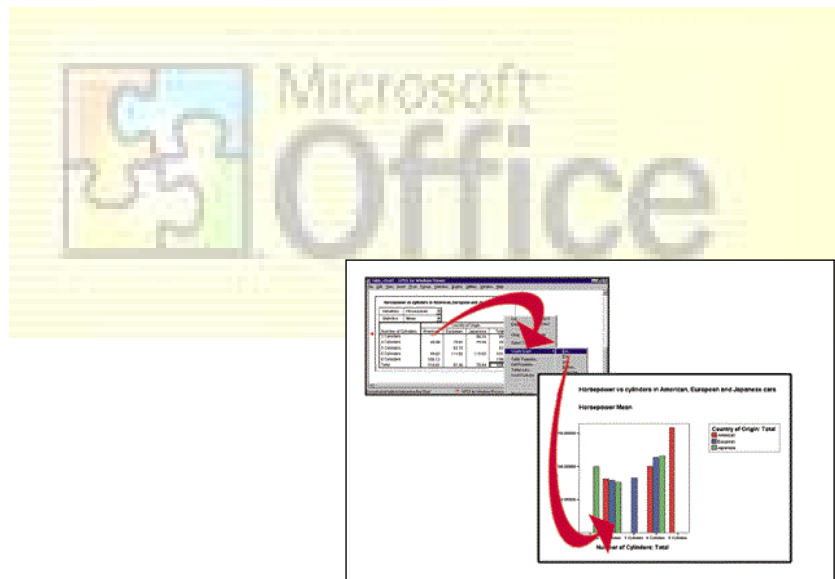


Computer skills for Statisticians Using MS Office and SPSS



Manual supporting the course "Computer Knowledge for STATISTICIANS"

compiled by Klaus Röder - march 2002/ revised may 2004

Table of Contents

1. MICROSOFT WORD.....	4
1.1. THEORETICAL BACKGROUND.....	4
1.2. THE PRACTICAL APPROACH.....	5
2. MICROSOFT EXCEL.....	21
2.1. THE PRACTICAL APPROACH.....	21
3. MICROSOFT POWERPOINT.....	30
3.1. INTRODUCTION.....	30
3.2. THE PRACTICAL APPROACH.....	30
3.3. THE GENERAL IDEA.....	33
3.4. ANIMATION.....	35
3.5. COMBINING SEVERAL SLIDES.....	36
3.6. TEXT EDITING.....	37
3.7. DRAWING.....	37
3.8. PACK AND GO.....	37
4. MICROSOFT ACCESS.....	39
4.1. DATABASES: WHAT THEY ARE AND HOW DO THEY WORK.....	39
4.2. TABLES: WHAT THEY ARE AND HOW DO THEY WORK.....	43
4.3. QUERIES: WHAT THEY ARE AND HOW DO THEY WORK.....	45
4.4. FORMS: WHAT THEY ARE AND HOW THEY WORK.....	46
4.5. REPORTS: WHAT THEY ARE AND HOW THEY WORK.....	48
4.6. DATA ACCESS PAGES: WHAT THEY ARE AND HOW THEY WORK.....	50
4.7. MACROS: WHAT THEY ARE AND HOW THEY WORK.....	50
4.8. MODULES: WHAT THEY ARE AND HOW THEY WORK.....	50
4.9. RELATIONSHIPS IN A DATABASE.....	50
5. HANDS ON SPSS.....	51
5.1. A GENERAL VIEW.....	51
5.2. HOW TO START AND STOP SPSS.....	51
5.3. SPSS WINDOWS.....	51
5.4. EXAMPLE: THE QUESTIONNAIRE OF THE MUNICH CENTRE'S SAMPLE SURVEY.....	52
5.5. THE ANALYSIS OF FREQUENCY.....	58
5.6. THE TRANSFORMATION OF VARIABLES.....	62
5.7. DEFINING VARIABLES.....	62
5.8. VARIABLE NAMES IN SPSS.....	62
5.9. VARIABLE TYPES.....	62
5.10. MISSING VALUES.....	63
5.11. VARIABLE LABELS.....	63
5.12. COMPUTE VARIABLES.....	65
5.13. RECODE VARIABLES.....	66
5.14. TRANSPOSE VARIABLES.....	67
5.15. SORT CASES.....	67
5.16. SELECT CASES.....	67
5.17. THE TRANSFORMATION OF DATA FILES.....	72
5.18. MERGING FILES: GENERAL CONSIDERATIONS.....	72
5.19. MERGE FILES: ADD CASES.....	72
5.20. MERGE FILES: ADD VARIABLES.....	73
5.21. FUNCTIONS FOR THE INITIAL ANALYSIS OF DATA.....	78
5.22. FREQUENCIES.....	78
5.23. EXPLORATIVE DATA ANALYSIS.....	84
5.24. CROSSTABS.....	88
5.25. SYNTAX FILES.....	93
5.26. SYNTAX FILES: GENERAL CONSIDERATIONS.....	93
5.27. EDITING TEXT IN OUTPUT AND SYNTAX WINDOWS.....	94

5.28.	EXECUTE COMMANDS IN SYNTAX WINDOWS	95
5.29.	THE COMMAND SYNTAX	95
5.30.	THE PRESENTATION OF DATA	98
5.31.	LISTINGS AND REPORTS OF DATA.....	98
5.31.1.	<i>List Cases</i>	98
5.31.2.	<i>Report Summaries in Rows</i>	99
5.31.3.	<i>Report Summaries in Columns</i>	100
5.32.	THE TABLES MODULE	104
5.32.1.	<i>Basic Tables</i>	105
5.32.2.	<i>General Tables</i>	107
5.32.3.	<i>Tables of Frequencies</i>	110
5.33.	GRAPHICS AND CHARTS.....	112
5.33.1.	<i>Graphs from the Graphs Menu</i>	113
5.33.2.	<i>Graphs resulting from procedure calls</i>	114
5.33.3.	<i>Graphs from the Interactive Graphs Menu</i>	114
5.33.4.	<i>Using the Chart Editor</i>	114
5.33.5.	<i>Interactive Graphs and the Interactive Chart Editor</i>	116


1. Microsoft WORD

1.1. Theoretical Background

1. INTRODUCTION:

Word is a WhatYouSeeIsWhatYouGet-text processor meaning that the resulting printout is permanently previewed, so you type your text on a virtual sheet of paper. This enables you to create text in a similar way as using an ordinary typewriter – which is often thought to be the easiest way. However, by doing so, you waste the advantages of word processors with regard to productivity and efficiency. So, the cause of this class is to enable you to work with WORD in an easy and efficient way. This will be done by defining paragraph styles for all kinds of paragraph styles that are used frequently in a document like headings, body text or footnotes

1.1. The GRAPHICAL USER INTERFACE of WORD.

Like all WINDOWS-based programs WORD is operated predominantly by using the graphical user interface, which in general means the use of menu-entries and command-buttons, the latter act as instant command execution with default setting for frequently used commands. E.g. by pressing  the actual document will be printed to the default printer without any dialogue. Menus, however, give access to nearly all commands and setting options (a small number is only available through the macro-language) and due to their amount are organized in logically combined submenus structures.

1.2. WORD PROCESSING CAPABILITIES: What can be done?

1.2.1. Letters

The probably most common professional use of WORD is for correspondenc. You will find a variety of templates to create personal or business letters by choosing. Since these are quite short documents the main advantage of the use of paragraph styles does not lie that much in a structured set-up within a letter but in the uniformity of all the letters which are based on a single template e.g. for corporate identity reasons.

1.2.2. Thesis/Publications

For these long documents it is essential to stick to the structuring by the use of paragraph styles in order to assure

- uniformity within the document
- reformat every paragraph based on a certain style simply by changing that style
- make use of automatic function like automatic generation of a table of content.

1.2.3. Books

See Thesis/Publications plus automatic header entries

1.3. WP RESTRICTIONS: What cannot be done resp. what should be done with another application.

1.3.1. Layout intensive publications (flyers).

1.3.2. Documents with lots of imported graphics (magazines).

1.3.3. Documents which size exceed 55,87cm × 55,87cm (posters).

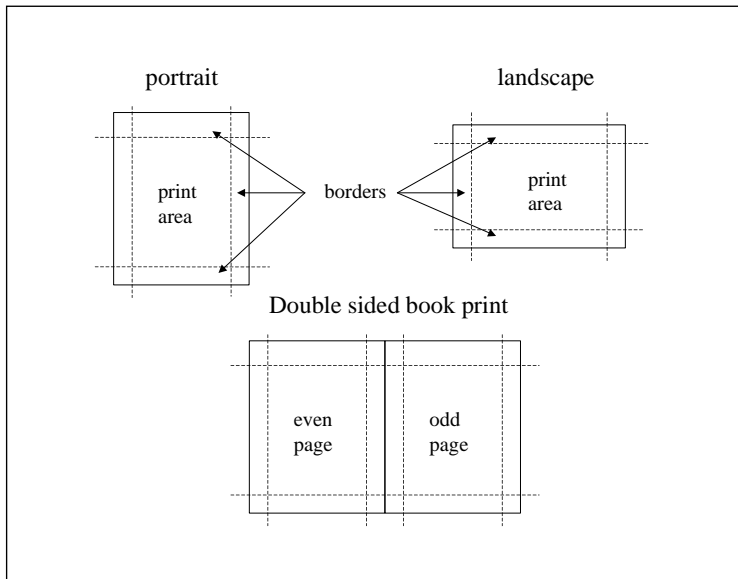
1.2. The Practical Approach

In general all documents should be structured in a way easily to understand. Concerning the layout this means: Do not overload a document with visual effects and eye catchers. Use a maximum of three fonts, better only two: A non-serif like arial for big headings and serif font for the rest esp. body text to enhance legibility. Another important suggestion is to keep in mind the purpose of one's writing: E.g. if the document is going to be copied you should use only black&white figures – avoid even grayscale.

2. PAGE SETTING:

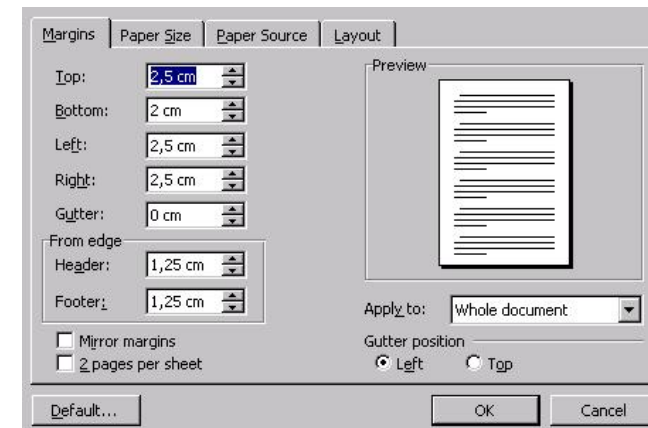
Generally described PAGE SETTING serve to define the paper size and the printing area within the paper size (see Figure 1-1) and the paper feeder.

Figure 1-1: Page Setting



To adjust the PAGE SETUP call the respective command within the FILE-menu. WORD shows a dialogue box with four registers:

Figure 1-2: Page Settings Dialogue

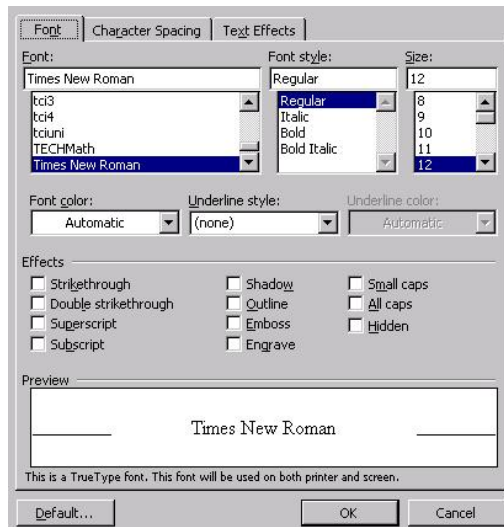


While PAPER SOURCE refers to the printer's feeder only you can use the remaining tabs to define the layout to meet your needs.

3. TEXTUAL ITEMS:

3.1. EDITING TEXT: Generally text is inserted at the insertion mark. This cursor can be positioned everywhere within an existing text but not beyond the last paragraph mark. You can select text (like every element of a document) either by moving the mouse pointer over the desired area while keeping the left mouse button down or by moving the insertion mark with the arrow keys while pressing SHIFT. Highlighted elements can be edited by copy and paste, drag & drop and formatted using the FORMAT/FONT-menu. You can see the most important options in Figure 1-3.

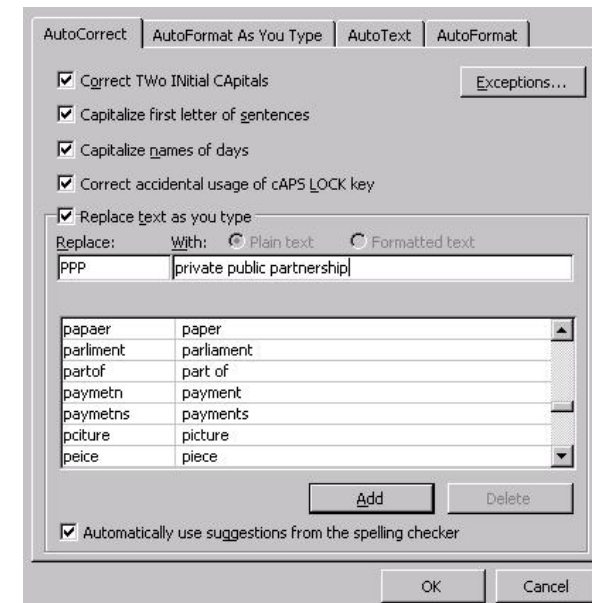
Figure 1-3: Format/Font



(Tab CHARACTER SPACING and TEXT EFFECTS refers to special effects only).

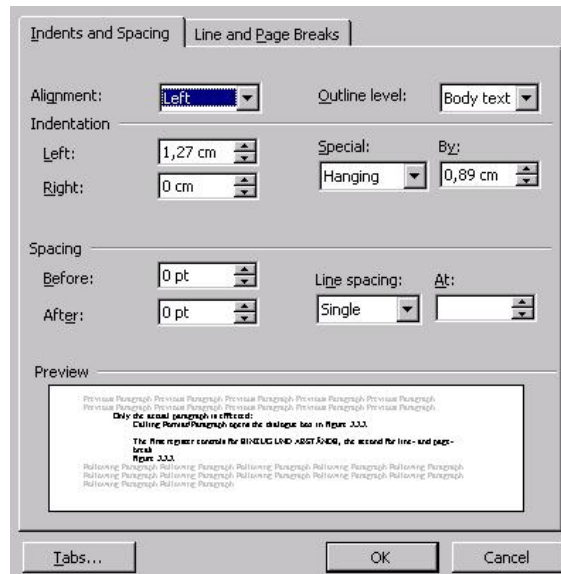
- 3.1.1. AUTO-CORRECTION: In order to edit frequently used subject-specific terms (e.g. public private partnership) by typing a custom defined abbreviation (e.g. PPP), just call AUTO-CORRECTION form the EXTRA-menu, where you can define abbreviation and substitute as shown in Figure 1-4.

Figure 1-4: Autocorrection



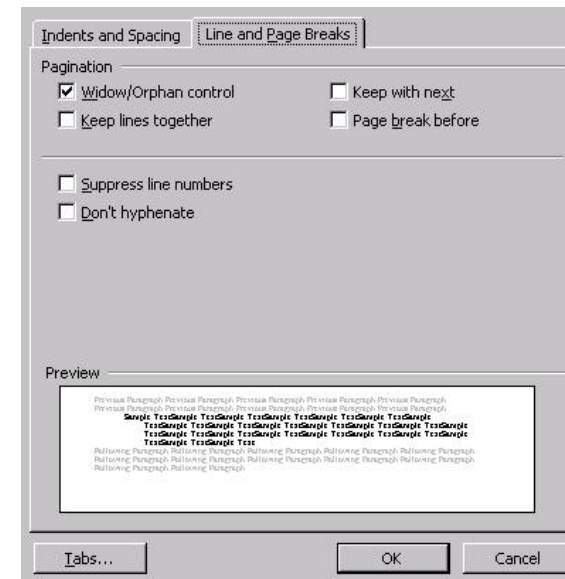
- 3.1.2. NON-BREAKING SPACE/HYPHENATION: Under certain circumstances you should suppress a break at the end of the line. By default, for example the expression "E-Mail" would be broken at the end of a line. By using CONTROL + SHIFT + DASH "E-Mail" the break is suppressed. With CONTROL + SHIFT + SPACEBAR you get the same effect for a non-breaking space.
- 3.2. GENERAL REMARKS: A paragraph can be defined as the content preceding a paragraph marks (¶), irrespective of that content's format (text, picture, empty space etc.). To mark a paragraph it is sufficient to place the insertion mark into the desired paragraph. The proper use of paragraph styles is essential with regard to efficient document creating. Later modifications of certain style are effective to all paragraphs, which are based on the appropriate paragraph. Additionally many functions in WORD are only available if the text is structured by the use of paragraph styles. You can choose an existing style from the style list or by FORMAT/STYLE. In this case choose a style out of the list on the left side and click APPLY.
- A paragraph can be edited in two ways:
- 3.2.1. Only the actual paragraph is effected:
Calling FORMAT/PARAGRAPH opens the dialogue box in Figure 1-5.

Figure 1-5: Format/Paragraph



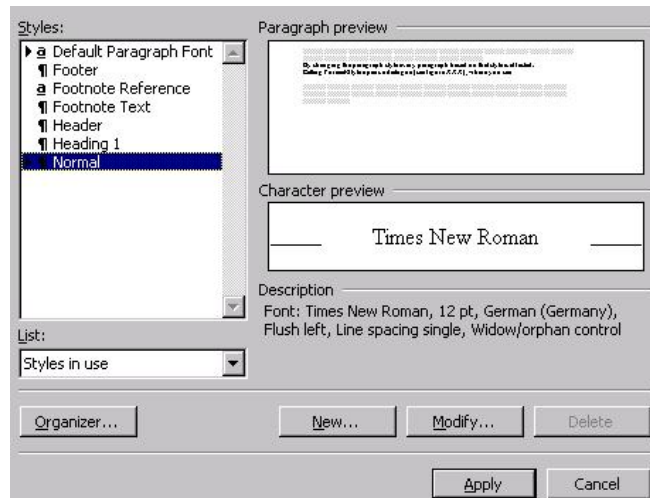
The first register controls for INDENTS AND SPACING, the second for LINE- AND PAGE BREAKS:

Figure 1-6: Line and Page Breaks



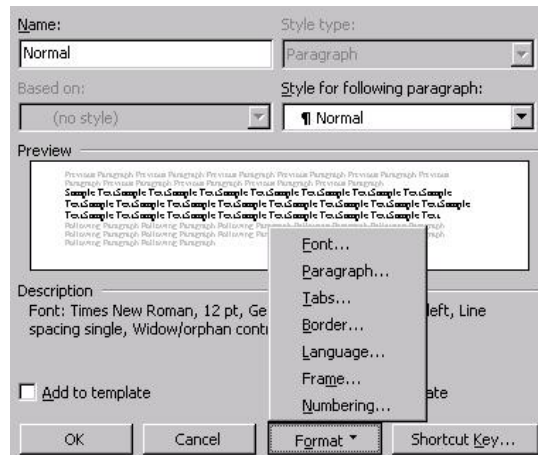
3.2.2. By changing the paragraph style every paragraph based on that style is effected.
Calling FORMAT/STYLE opens a dialogue (see Figure 1-7).

Figure 1-7: Format/Style



There you can choose a style and then modify it with a click on the appropriate command button, that opens the following dialogue and gives access to a variety of format options via the FORMAT-button as shown in Figure 1-8.

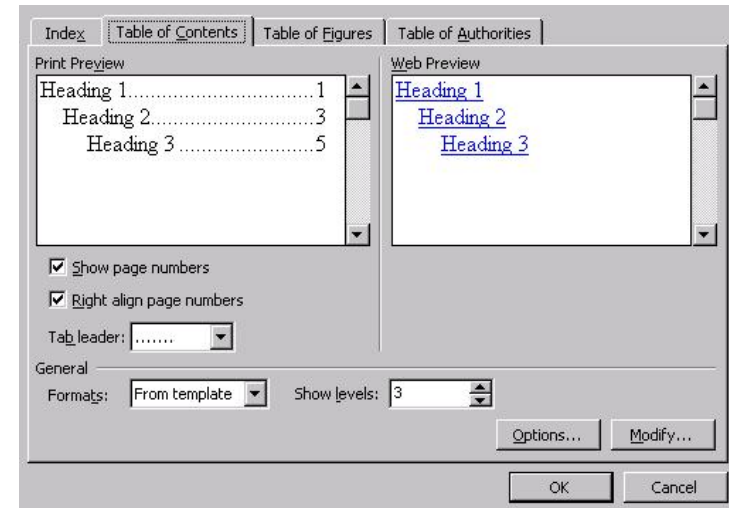
Figure 1-8: Style, Format Options



3.3. HEADINGS: You will structure long documents by using headings to divide chapters and sub-chapters. Formatting these headings with specified paragraph styles for each

level of headings enables WORD to distinguish headings from other elements of the text and thus create a table of contents automatically: Just place the insertion mark at the position where you want to create the table of content and call INSERT/INDEX AND TABLES. In the opening dialogue box you choose TABLE OF CONTENTS (see Figure 1-9).

Figure 1-9: TOC



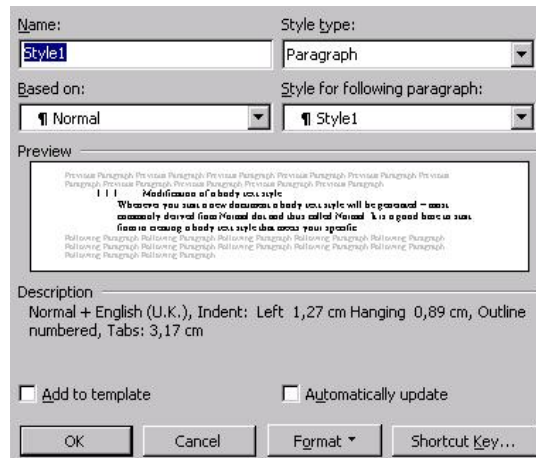
Here you may adjust the layout according to your preferences.

3.4. BODY TEXT:

3.4.1. Modification of a body text style.

Whenever you start a new document a body text style will be generated – most commonly derived from NORMAL.DOT and thus called NORMAL. It is a good base to start from in creating a body text style that meets your specific needs best. To create a style named BODY TEXT based on STANDARD you call FORMAT/STYLE (see Figure 1-7) and click NEW. You will see a dialogue box like Figure 1-10, where you replace the default name STYLE1 with BODY TEXT.

Figure 1-10: Style definition



Make sure that drop-down list BASED ON the style NORMAL is displayed. Then click the FORMAT-button and choose PARAGRAPH. You will be shown the ordinary paragraph dialog box where you can adjust your settings. Then you confirm all dialogues with OK. The new style BODY TEXT is now available in the above described way (see 3.2).

3.4.2. Creation of a MACRO

There are two ways of creating macros: You can either edit a macro in the macro-program-language (WORD uses VISUAL BASIC FOR APPLICATION) or you can record a macro. We will restrict ourselves to the latter since this does not imply any programming skills.

To record a macro it is helpful to assist for repetitive jobs e.g. if you frequently make use of a certain command that is "hidden" deep in sub-menus. Additionally every time you call the commands you may have to go through a lot of settings and so on.

Imagine you edit an English text with a lot of French quotations. If you just type in the French parts, the spell checker will stop at every single French word. To avoid this, it is reasonable to change spell-checking to French for the quotation. This is the usual way to do so:

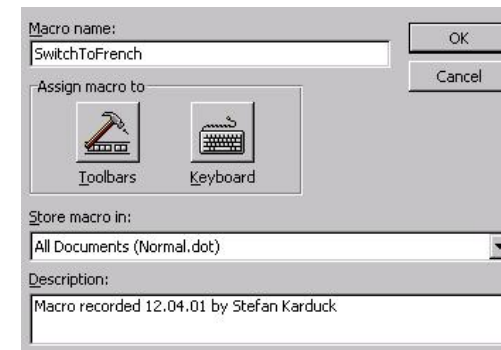
Select the desired section. Call TOOLS/LANGUAGE/SET LANGUAGE..., go through the displayed list of available languages in order to choose FRENCH. Then confirm with OK.

Figure 1-11: Language selection



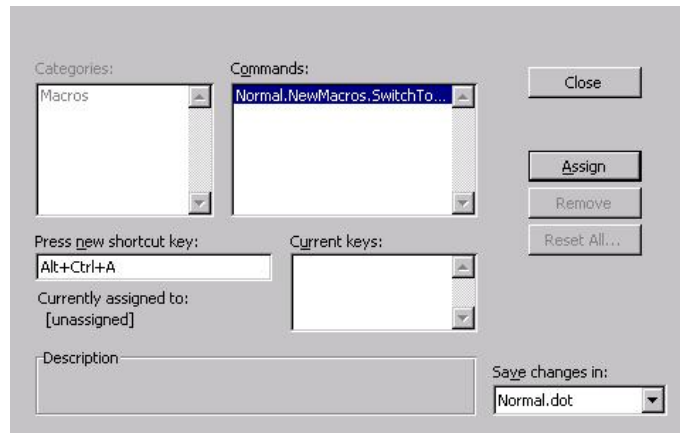
Without the help of a macro it is necessary to repeat all these steps for every single French quotation. To automate switching the spell checker you call TOOLS/MARCOS/RECORD NEW MACRO. The following dialogue box will appear:



Figure 1-12: Macro Start-up Box



Type in a name (e.g. SwitchToFrench) for the new macro, then click the KEYBOARD-button and assign a key-stroke (e.g. CONTROL + ALT+ A) in the next dialogue (see Figure 1-13 below);

Figure 1-13: Assign to keyboard



WORD checks whether a proposed combination is used already, so that you may assign an alternative one. Click ZUORDNEN. After closing this dialogue with CLOSE you will find a cassette tape icon added to the mouse pointer to indicate that recordings are going on and the additional buttons  are displayed. Now you may repeat the above described steps to switch the spell checker, except selecting anything. After finishing click on  to end the recording. Since the macro is available immediately you can switch the spell checker settings for a selected part of text with the assigned keystroke – ready.

4. NON-TEXTUAL ITEMS:

- 4.1. **Figures:** You can import figures via copy and paste from other applications or via INSERT/PICTURE/FROM FILE. This command starts the ordinary OPEN-dialogue, where you can search and choose the graphic-file to be imported. In order to label and to number graphics consecutively select a graphic and call INSERT/CAPTION.

Figure 1-14: Insert/Caption

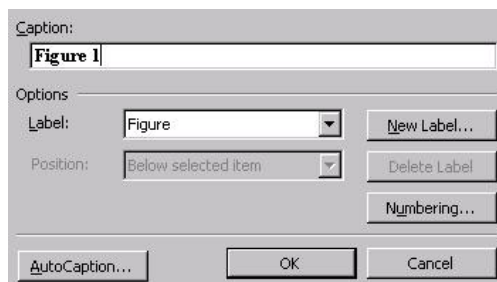


Figure 1-14 shows the dialogue where you choose the type of label, the labels position

and the label text. Confirm with OK and your label will be displayed at the desired position.





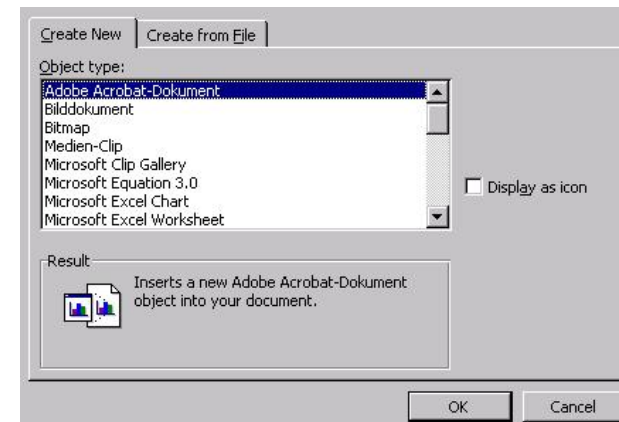
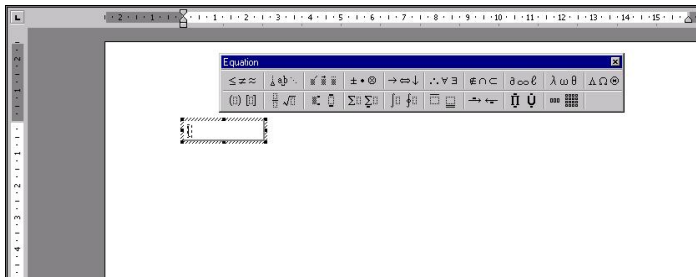
- 4.2. **TABLES:** WORD gives you lots of possibilities to create and import tables of which the two most important shall be discussed now: Position the insertion mark at the position where a new table shall be created. A click on  opens a symbolic table that allows you to choose the number of rows and columns you need just by moving the mouse pointer over that symbolic table. With a click on the left mouse button you confirm the range and Word creates the corresponding tables at the cursor's position. With  you can draw additional cell into such a table. Using  instead of  works just the same but generates an EXCEL-table within WORD so that you have access to all spreadsheet-functions. The labelling can be done just as shown for graphics.
- 4.3. **EQUATIONS:** WORD itself is not able to generate equations but it can handle them as embedded objects. That means equations must be provided by some kind of editor. Since there is an editor installed along with most MS-Office configurations you should at least be able to use MICROSOFT EQUATION. Then you call this editor by INSERT/OBJECTS. A dialogue like Figure 1-15 will be displayed so that you can choose the equation generating application (the exact content of those lists depends on your system's configuration).

Figure 1-15: Insert/Objects



This will open the editor and a position frame as Figure 1-16 shows. You can use the tool bar to insert elements of the equation that are displayed inside the position frame at once.

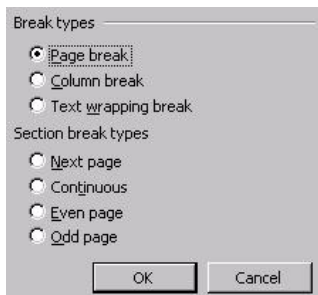
Figure 1-16: Equation Editor



After editing the equation you can return to WORD with a mouse click on the actual document. This automatically imports the new equation that now can be labelled in known manner.

5. **HEADER:** To toggle normal and header view call VIEW/HEADER. Headers are defined throughout a section, so that their content will be repeated on every page of a section. To insert a new section choose INSERT/BREAK and choose an option out of the list SECTION BREAK TYPES.

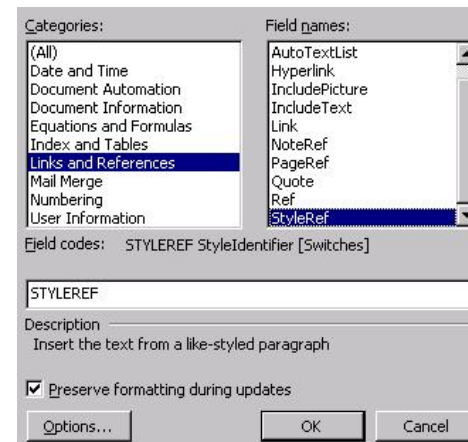
Figure 1-17: Breaks



In order to allow headers to become dynamic, e.g. to display the different contents of level II-headings you can define reference fields instead of fixed entries.

This must be done with the cursor inside the appropriate header (in "header-mode"). Then call INSERT/FIELD and choose StyleRef out of Links and References-categories as shown in Figure 1-18.

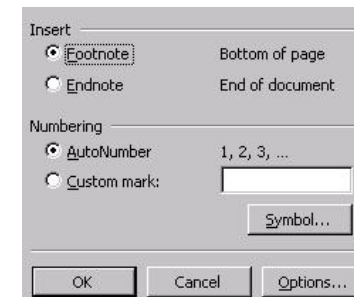
Figure 1-18: Insert/Fields



Click the OPTIONS...-button to open the STYLES-tab where you can choose a heading style (here: level) that should serve the content for the header.

6. **FOOTNOTES:** You insert a footnote at the actual position of the insertion mark by calling INSERT/FOOTNOTE. The default setting of the dialogue box shown in Figure 1-19

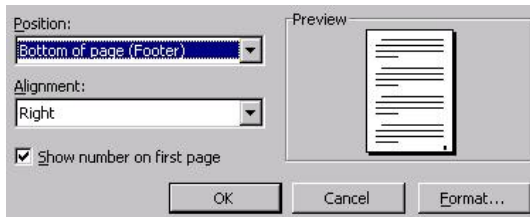
Figure 1-19: Insert/Footnote



can be accepted with Ok. Word inserts the footnote marker and moves to the end of the page, where a section for the footnote text is provided. Since footnotes are handled just like ordinary paragraphs (after inserting the first footnote, you will find the automatically generated footnote style in the STYLE-list. The only thing to remember is: You cannot delete a footnote just by killing the footnote text. Instead you must delete the footnote marker!

7. **PAGE NUMBERING:** Choose INSERT/PAGE NUMBERS. The appearing dialog box (see Figure 1-20) prompts for basic setting like position and justification.

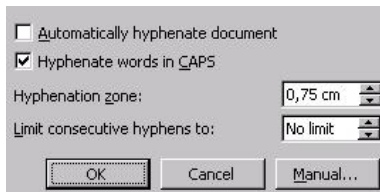
Figure 1-20: Page numbering



If you need more options click the appropriate button.

8. **HYPHENATION:** WORD provides an automatic hyphenation option. Unfortunately this hyphenation is not free of mistakes so that it is advisable to choose the discretionary hyphenation. To do so call **TOOLS/LANGUAGE/HYPHENATION**.

Figure 1-21: Hyphenation



Click **MANUAL**. WORD starts to work through the text proposing possible hyphenations, that you have to accept or reject.


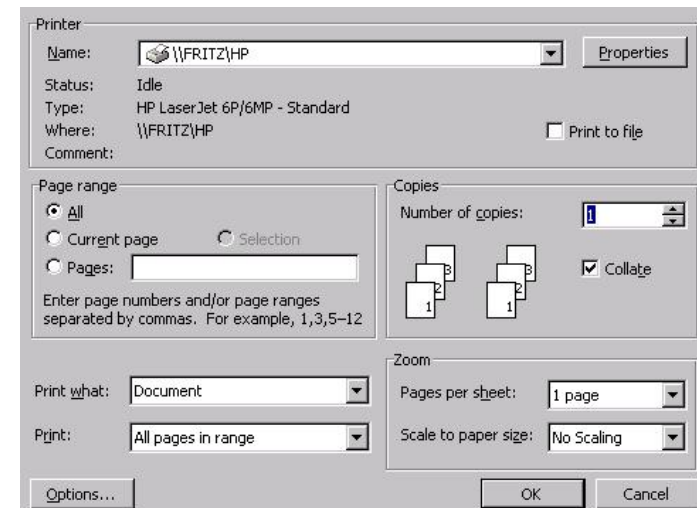
9. **SPELL CHECKING:** Since WORD 7.0 this processor offers background spell checking marking misspells with a red wavy line e.g. mistak for misspellings and a green wavy line for grammar mistakes. So you may step through the text in order to control for these marks. Or you call **SPELLING AND GRAMMAR** from the **TOOLS**-menu. A dialogue will pop up and guide you through the process.
10. **PRINTING:**
 - 10.1. **HARDCOPIES:** The easiest way to print a document is to click . The document will be sent with default setting to the default printer. With **FILE/PRINT** you gain control over all possible options as Figure 1-22 shows.

Figure 1-22: Print



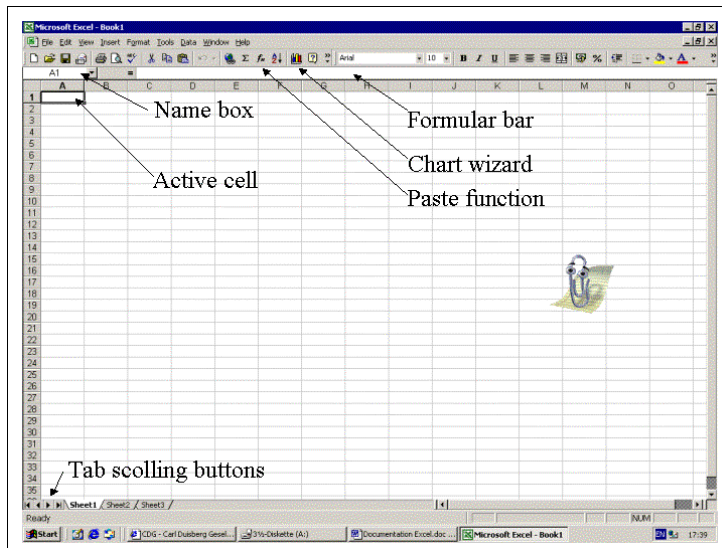
- 10.2. **Printing to a POSTSCRIPT FILE/PDF-FILE for distribution:** To distribute your documents electronically it is advisable not to choose WORD's own *.doc-format. First, this format is quite memory-intensive, second, docs generated with an actual version of WORD cannot be read with older versions, third, your layout will change more or less from machine to machine depending on the printer driver used and fourth, your documents are open to manipulation. A very elegant way to avoid these disadvantages is to print your document to a portable file-format. Most popular are Adobe's PostScript for high ends printouts and Portable Document File (PDF) for electronic distribution. While you can find PostScript-drivers on every WINDOWS-CD for free, you need to buy a PDF-driver for direct pdf-generation. However, since pdf is a subset of the PostScript-specifications there are programs available in the web for free that can convert PostScript to pdf. Best known is Aladdin's GhostScript that can be downloaded from <http://www.aladdin.com/>.


2. Microsoft EXCEL

2.1. The Practical Approach


10.3. The GRAPHICAL USER INTERFACE OF EXCEL.

Figure 2-1: GUI



The main idea of controlling EXCEL by menu commands as well as by command buttons is just the same as seen in WORD already. They simply differ in the commands available, e.g.  to create charts is not available in WORD. But there are still some other differences in layout:

The working area of EXCEL is made up from cells. The actual active cell(s) are marked with a frame. These coordinates within the spreadsheet can be seen in a drop-down-list in a display right above the spreadsheet (see Figure 2-1) selected; right hand, following “=” you see the formula used in that cell, which is quite important: Imagine a cell displaying “25” in the spreadsheet. This might be, because somebody types in “25” – or it might be the result of a calculation, lets say “=5*5”. You can check it by selecting that cell.

Another important thing to remember is, that an EXCEL-file can contain several spreadsheets, that can be interactively connected. You navigate through the sheets with  **Sheet1** / Sheet2 / Sheet3 / on the left bottom of the screen.

10.4. SPREADSHEET CAPABILITIES: What can be done?

10.4.1. Easy editing and presenting data.

10.4.2. Calculation with built-in formulas.

EXCEL delivers a wide variety of built in functions. You can simply type them into a cell as a command or access them with the help of a wizard. Regardless of your approach, all kinds of calculations must start with an “=”.

10.4.3. Creation of graphical presentation of data.

This is surely one of the applications highlights. With the help of a wizard you can create meaningful graphics at a glance.

10.4.4. Simple statistics

EXCEL allows statistical calculation based on raw data as well as on frequencies. The latter is not possible with most statistical packages.

10.5. SPREADSHEET RESTRICTIONS: What cannot be done resp. what should be done with another application.

10.5.1. Advanced statistics.

The use of EXCEL is not useful for statistics exceeding bivariate analysis.

10.5.2. Advanced graphical presentation of data.

EXCEL cannot display more than two dimensions (variables) at a time. So called 3d-charts are just for the visual effect.

10.5.3. Conditional calculation (e.g.: One column contains data about the sex and another column about age. It is not possible to calculate the average age of males only directly. Either it is necessary to resort the data or to edit a macro.)

11. EDITING DATA:

11.1. Spreadsheet layout

A general advise is to layout a spreadsheet like this:

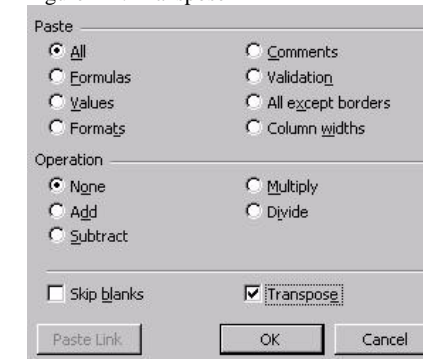
1 case = 1 row, 1 variable = 1 column.

This ensures easy exports to statistical applications like SPSS which in general stick to this layout.

11.2. Transposing.

If – for any reason – the layout is found right the other way around, you can transpose the data: Copy the desired range to the clipboard, switch to a new sheet and call PASTE SPECIAL from the EDIT-menu. EXCEL displays the box in Figure 2-2

Figure 2-2: Transpose



where you check TRANSPOSE and confirm the remaining settings with OK.

11.3. Autofill.

This is quite an important feature to make spreadsheet calculations faster: Imagine you have to create a new data sheet. The first column is to contain the case numbers – lets say up to 250 cases. Typing them case by case is boring and inefficient. With AUTOFILL you simply type in the first two cases in neighbouring cells in order to “indicate the system” and select them. Then you move the mouse over the small cross at the right bottom of the selection (see below).

Case
1
2

After that keep the left mouse button down and draw down. EXCEL fills in the numbers automatically and displays the progress in a yellow pop-up.

But this does not work only with numbers: EXCEL will “understand” more or less every common repeating system e.g. months or days of the week.

11.4. Importing data.

By default EXCEL supports most spreadsheet-formats. They can be loaded with the ordinary OPEN-dialogue. There you should specify the desired format in the lower drop-down-list. Additionally most data base applications feature an EXPORT TO EXCEL-option. The data will then be written to an EXCEL 4.0-file. However, in some cases you have to access a data base directly. With DATA/GET EXTERNAL DATA/NEW DATA QUERY you can specify the data format to be read manually.

12. CALCULATION:


12.1. Standard operations.

Select a cell and type in a formula starting with “=”.

12.2. Autocalculation.

Imagine you want to calculate the sum of several columns. For the first column you have to do it “by hand” then select that cell again, position the mouse pointer over the small cross and draw it to the right (or left) with left mouse button down.

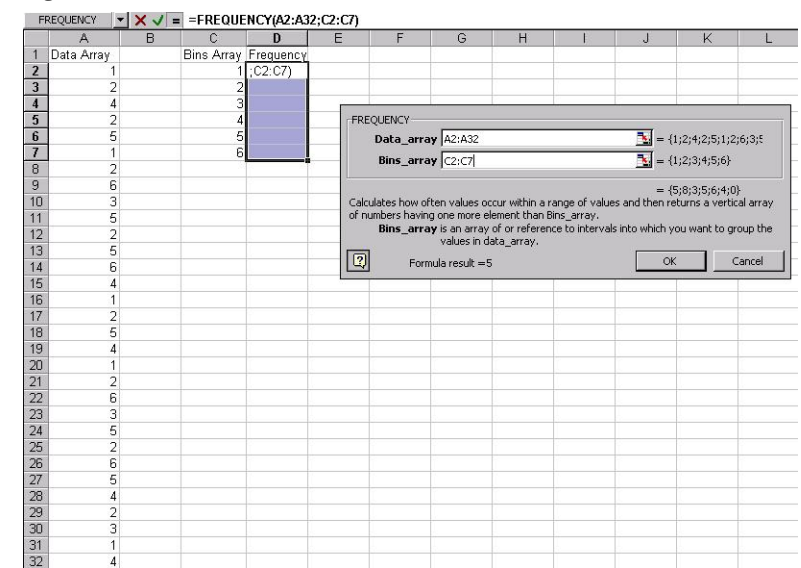
12.3. The use of simple formulas.

Simple formulas produce an output in just one cell. You can just type them or use the wizard with INSERT/FUNCTION or .

12.4. The use of matrix formulas.

In some cases, formulas deliver outputs to multiple cells. E.g. if you want to calculate a frequency matrix as figure shows:

Figure 2-3: Matrix Formula



In column A you find the data array with data varying between 1 and 6, in column C you give Excel the appropriate bins array. Then you highlight the cells where the formula output will be delivered to (here: D2:D7) and start the formula wizard to select FREQUENCY out of the STATISTICS-category. Excel asks where the data is to be found (see figure).

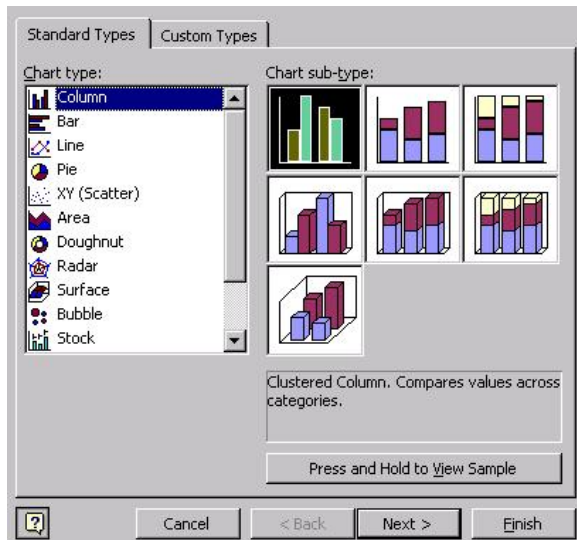
But instead of confirming your inputs with OK you have to press SHIFT + CONTROL + RETURN to point out that you want to multiple outputs.

13. GRAPHIC PRESENTATION:

13.1. Introduction to various types of charts.

The type of chart you choose should fit to the data and its scale quality. By default EXCEL offers a wide variety of charts as can be seen in Figure 2-4:

Figure 2-4: Chart Wizard 1 of 4



13.2. Creation of charts.


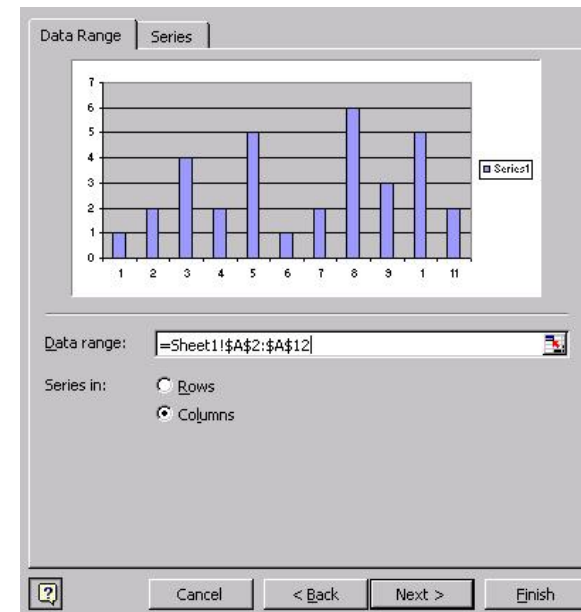
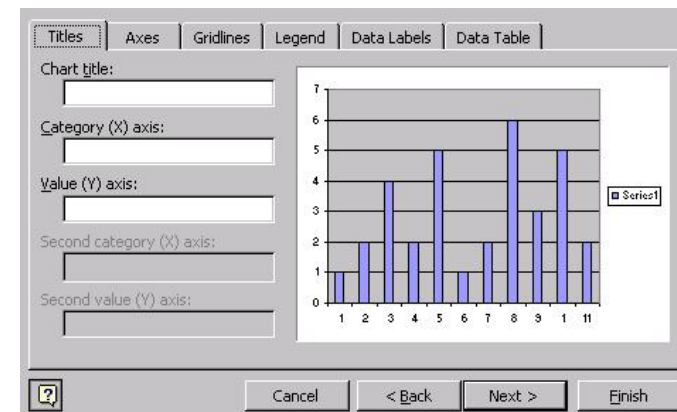
To create charts select the range of data and click . You will be guided through the process by a wizard. The first step is to choose the type of chart as can be seen in Figure 2-4. After you finished your choice please go on with a click on NEXT. The next dialogue serves to define the data range or – in the series tab – to define series of data ranges. After selecting some data, EXCEL displays with a draft preview of the chart.

Figure 2-5: Chart Wizard 2 of 4



With another click on NEXT you get access to a set of layout option as shown in Figure 2-6.

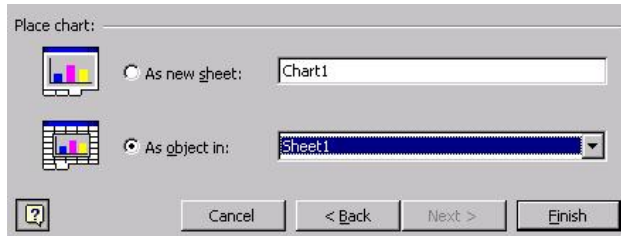
Figure 2-6: Chart Wizard 3 of 4



Again a click on NEXT allows you to choose whether the chart should be placed into

the active sheet or should be created as a new sheet.

Figure 2-7: Chart Wizard 4 of 4

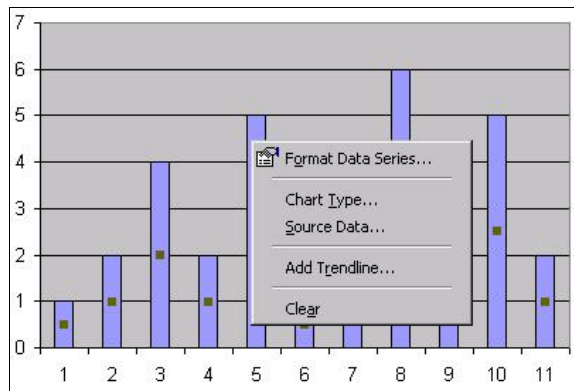


With FINISH the chart will be created.

13.3. Editing charts.

To edit an existing chart select the element to be edited and click it with the right mouse button. A context menu appears that displays all available options for the selected element (see Figure 2-8 as an example for a context menu related to right clicks on the columns).

Figure 2-8: Editing Charts



14. STATISTICS:

- 14.1. Frequencies, calculation and graphical presentation.
- 14.2. Mean and standard deviation.
- 14.3. Different types of regression models.
- 14.4. Smoothing a time-series and graphical presentation.

15. PRINTING:

- 15.1. Preview.

Since EXCEL does not display the page layout by default it is reasonable to preview a spreadsheet before printing with . EXCEL displays the layout with standard settings

that can be changed by pushing the LAYOUT-button on top of the screen. Please note the possibility to scale the output which is helpful if you want to place a huge range of data on a single sheet of paper.

16. COMBINING WORD and EXCEL to create a series letter/mail merge.

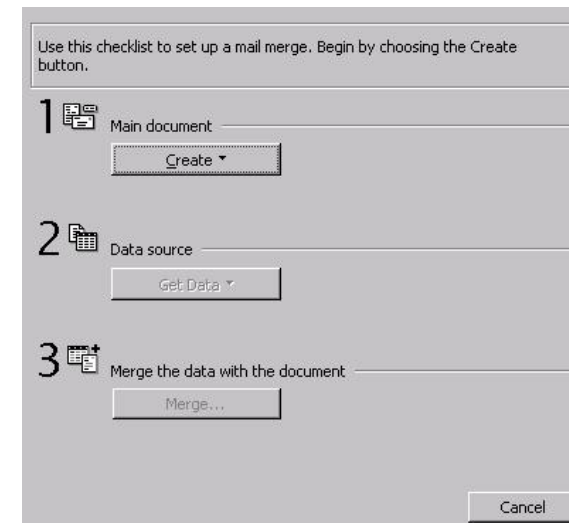
Series letters are used to produce form letters, mailing labels, envelopes, catalogues, and other types of merged documents. To do so EXCEL can serve as a container e.g. for address information as Figure 2-9: Data Source shows:

Figure 2-9: Data Source

	A	B	C	D
1	sa	surname	name	
2	Mister	James	Bond	
3	Miss	Pamela	Anderson	
4				

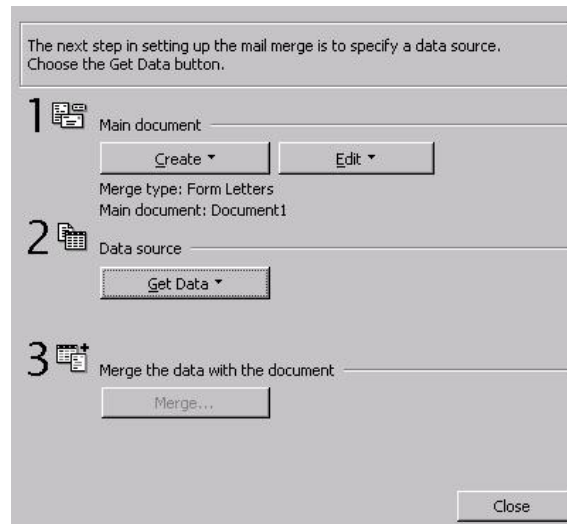
WORD will use the entries in row 1 as references to fields for the mailing labels within a series letter. After you have entered all labels to Excel save the file (here: as SeriesLetter.xls). Then start WORD with an empty document and call TOOLS/MAIL MERGER. This opens the dialog box in Figure 2-10:

Figure 2-10: Mail Merge Helper



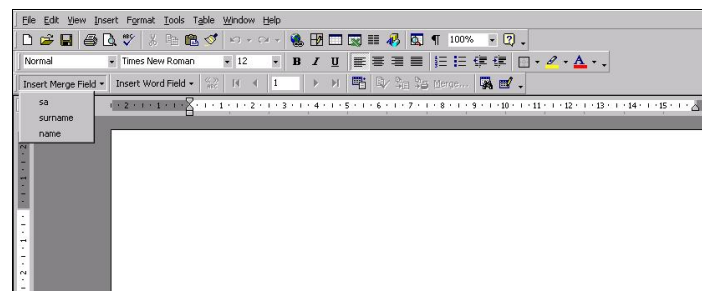
Click CREATE and choose FORM LETTER. WORD asks whether you want to create a new document or use the active document. Choose ACTIVE DOCUMENT. Then you need to establish the link to the address information stored in the prepared EXCEL-file. So click GET DATA.



Figure 2-11: Data linking



Choose OPEN DATA SOURCE. You will find a standard OPEN-dialogue to open SERIESLETTER.XLS. Please make sure to choose the appropriate type of file. Confirm to use the entire file and – after that – to edit the main document. You will find a new tool bar attached to the existing ones:

Figure 2-12: Mail Merge Tool Bar



With a click on INSERT MERGE FIELDS you can input fields according to the data source to the main document wherever you want. By default these fields display the fieldnames only. You toggle between fieldnames and field-contents with . And with  you can navigate through all cases contained in the data source file.

3. Microsoft POWERPOINT

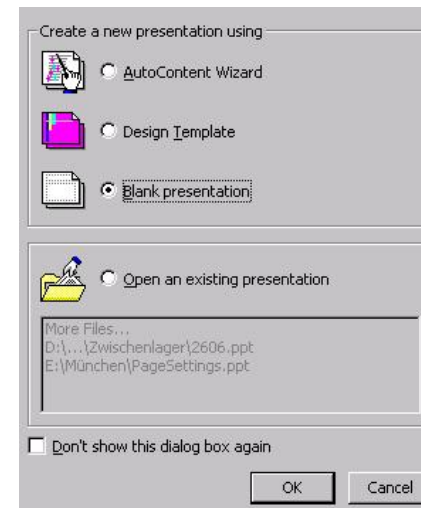
3.1. Introduction.

17. The GRAPHICAL USER INTERFACE of POWERPOINT.
18. CAPABILITIES: What can be done?
 - 18.1. Simple graphic editing.
 - 18.2. Combing lots of different file formats on one transparency: Graphic, text, animation, sound, video.
 - 18.3. Editing vector and bitmap graphics.
 - 18.4. Easy creation of graphic intensive presentation (for screen and real transparencies).
 - 18.5. Export a transparency to a variety of graphic formats.
 - 18.6. Export presentations to a PPT-viewer, so that presentations can be performed even on computers where PowerPoint is not available.
19. RESTRICTIONS: What cannot be done resp. what should be done with another application.
 - 19.1. Advanced graphic design.
 - 19.1.1. Presentations that are not sequentially arranged.

3.2. The Practical Approach

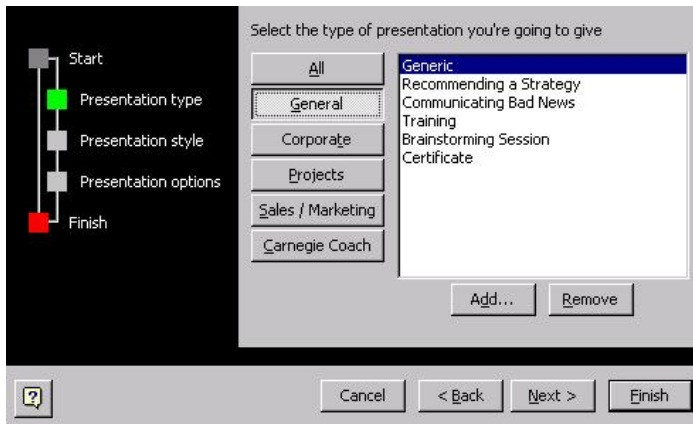
After starting POWERPOINT the program welcomes you by default with the dialogue box in Figure 3-1. Here you find the AUTOCONTENT-Wizard and that provides a good introduction to itself.

Figure 3-1: Start up box



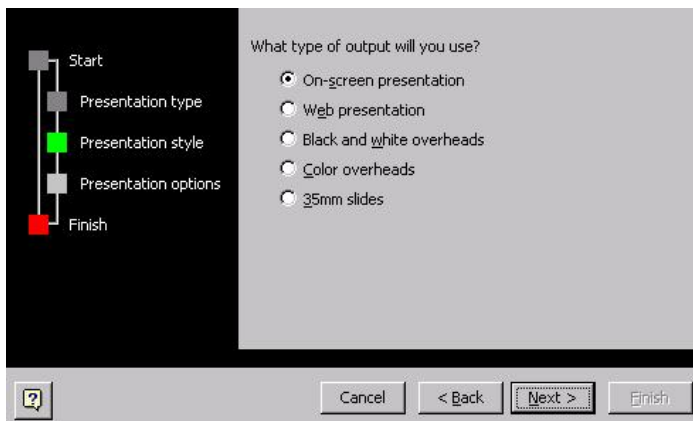
Start the wizard, choose the appropriate option and click OK and in the following dialogue NEXT (see Figure 3-2).

Figure 3-2: AutoContent Presentation Type



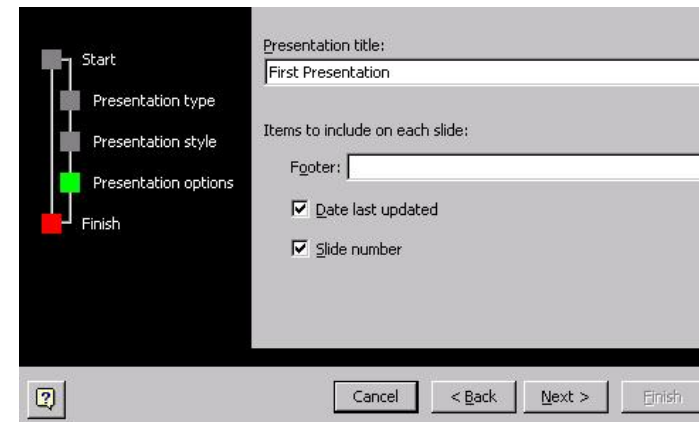
Here you get displayed the available templates according to the categories you choose with the command button in the middle of the box. For the explanatory example choose TRAINING out of GENERAL. To go on click NEXT again.

Figure 3-3: AutoContent Presentation Style



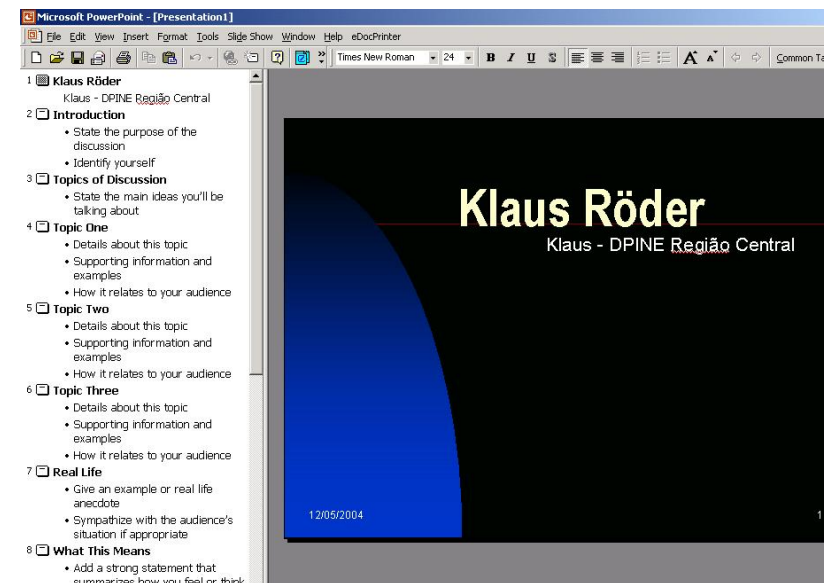
In Figure 3-3 choose ON SCREEN PRESENTATION from the style list and click NEXT to come to Figure 3-4.


Figure 3-4: Autocontent Presentation Options



Here you may type in a title for your presentation, a footer that will be repeated on all further slides. With a click on NEXT POWERPOINT confirms that all necessary inputs have been done, so click FINISH. With your inputs POWERPOINT has generated a ready made presentation (see Figure 3-5) that you can adopt to meet your needs.

Figure 3-5: First Presentation



With , that can be found in the left corner of your screen you can start the presentation. Mind the different transitions from slide to slide.

By playing with this presentation you can notice some reasonable hints:

- Do not fill up a slide too much.
- Spare on visual effects.
- Keep the background calm (you will find more background templates with FILE/NEW/DESIGN TEMPLATES)
 - Concerning the use of color.
 - Concerning the use of pictures etc.

3.3. The general idea

The general idea of POWERPOINT is to enable the user to combine more or less every sort of elements on a slide. During the presentation these elements can be displayed at once or step by step with special animation effects. To complete a presentation you can combine several slides to form a show.

All possible types of contents can be accessed via the INSERT menu. So I will mention the tool bar buttons only in the following. You can insert:



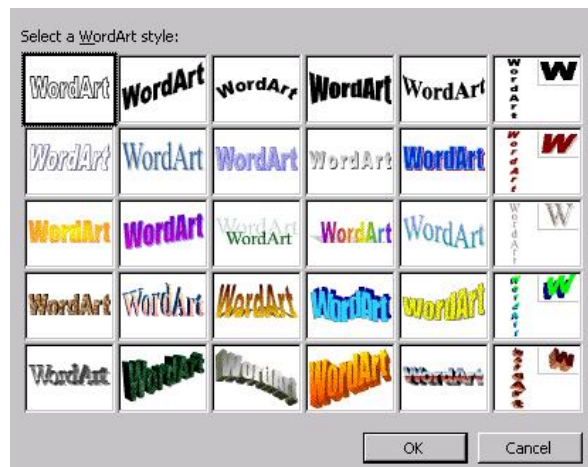
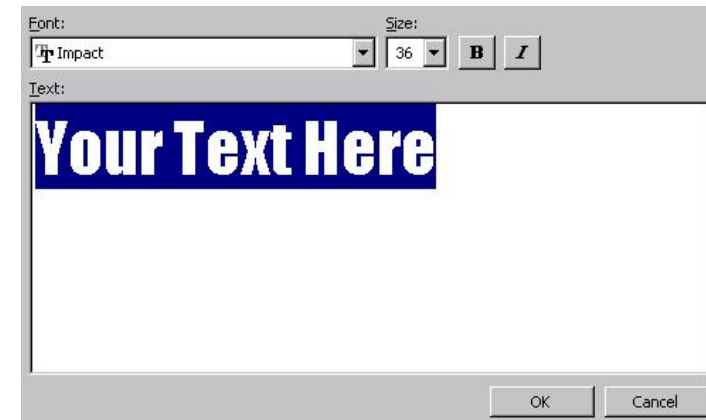
- Standard text: Text can be put on a slide inside of so-called textboxes by clicking . Then move the mouse pointer onto the slide to the desired position; you will see it changes its shape. Click once and POWERPOINT provides a textbox at the actual position of the mouse pointer. You can now insert your text; the box will enlarge automatically to fit the length of you typing.
- Word Art Text: You use this type of text to produce special text effects like 3d or text written in wavy lines and so on. With a click on , you get the Word Art Gallery displayed (see Figure 3-6).

Figure 3-6: Word Art Gallery



Here you select a given template that fits your conceivability best and click OK. The next dialogue (see Figure 7) prompts for the text to be written in Word Art.

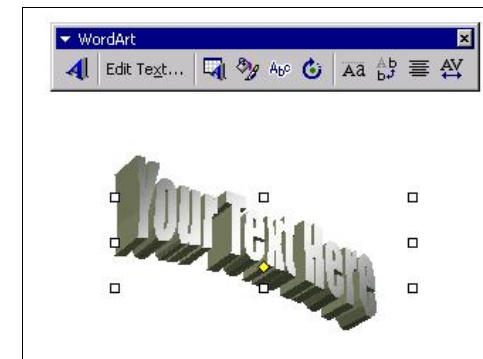
Figure 3-7: Word Art Text Editor



Beside this you can control for some fonts formats on top of the text entry field. After you have finished your text entry, POWERPOINT inserts the text into the actual slide with a click on OK.

As Figure 3-8 shows, the text is displayed together with a Word Art tool box and special crops mark around the Word Art text to edit and change the look of the Word Art Text.

Figure 3-8: Word Art Text




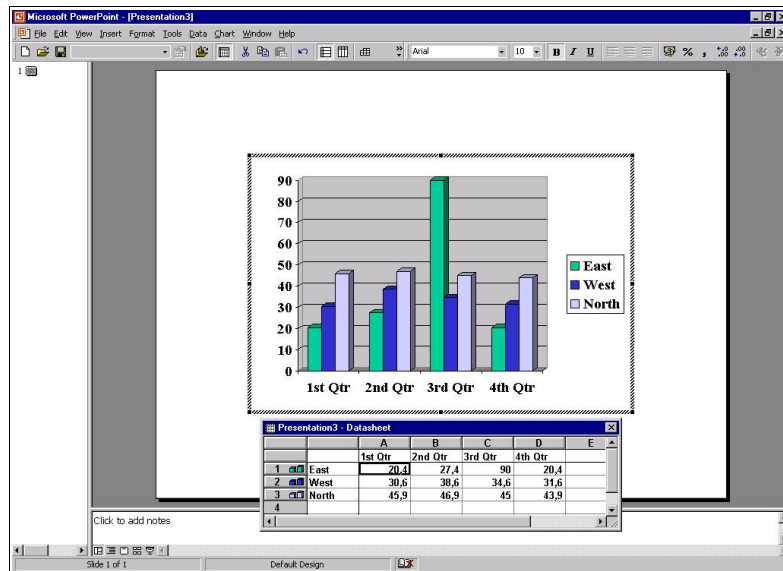
- Charts: With  you can insert charts to a slide. Differing from EXCEL, POWERPOINT does not start up with a chart wizard. Instead it provides a simple ready made chart and displays a data table.

Figure 3-9: Charts



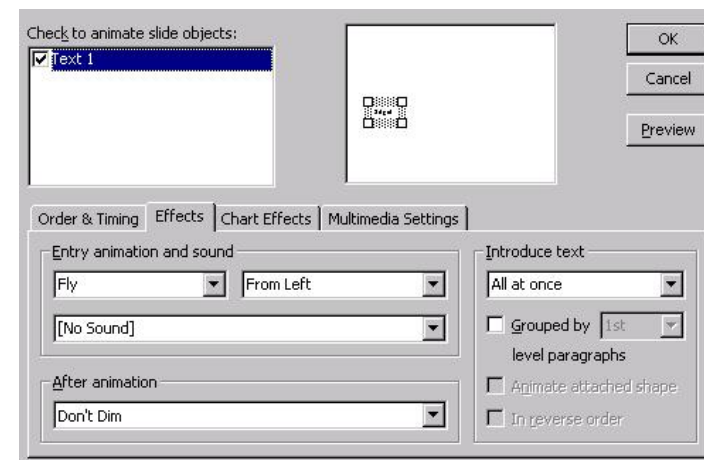
You can change this data to fit you needs and then change the chart just as you are used to in EXCEL. You can even change the type of chart.

- Tables: You can create and edit tables with . They can be handled just like in WORD.
- Movies and Sound: If your hard- and software configuration supports movies and sounds you can add them to a slide with the appropriate command from the INSERT menu.
- Pictures: With you can insert pictures from the Gallery. To use pictures from file call INSERT/PICTURE/FROM FILE.

3.4. Animation

All these items can be animated with special effects. To do so highlight the element to be animated with a click, open the SLIDE SHOW menu and call PRESET ANIMATION for a variety of ready made animation settings or CUSTOM ANIMATION for full control over the animation via four tab shown in Figure 3-10.

Figure 3-10: Custom Animation



3.5. Combining several Slides

To add another slide to your presentation call INSERT/NEW SLIDE. This new slide can be filled with new items and so on. During the slide show, these slides will be presented one after another. To enhance the presentation you can add slide transition effects to this procedure that works similar to the animation effects. Highlight a slide for which you want to define transition effects and choose SLIDE TRANSITION out of the SLIDE SHOW menu.

Figure 3-11: Slide Transition

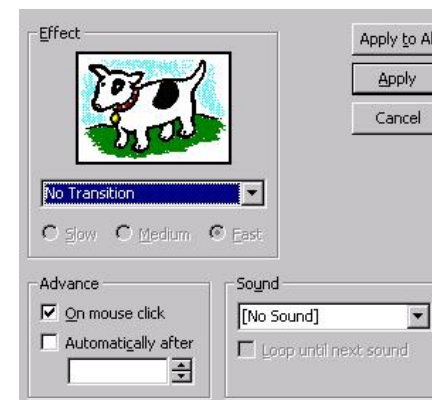


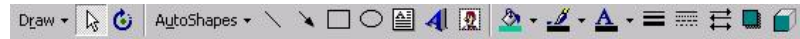
Figure 3-11 shows the appropriate dialogue and the options available.

3.6. Text editing

Beside the above described use of text boxes you can edit text into standard boxes by using the list on the left side of your screen (see for example Figure 3-1). Additionally you can use the key-strokes CTRL+ENTER to toggle between “new slide” or “new bullet list”.

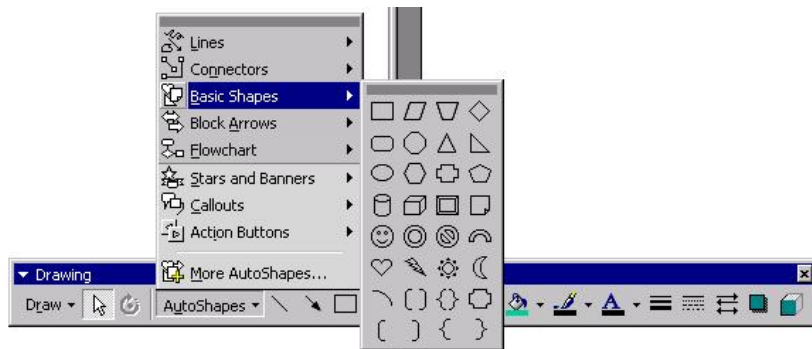
3.7. Drawing

With the tool bar



POWERPOINT offers limited drawing abilities. You will find the AUTOSHAPES-button especially useful for that purpose as Figure 3-12 shows.

Figure 3-12: AutoShapes



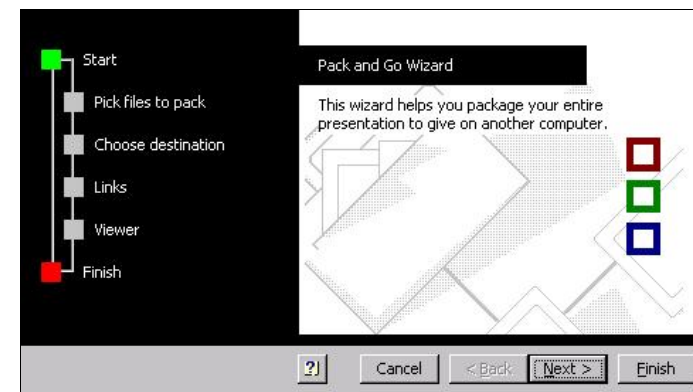
3.8. Pack and Go

In most cases you will not use a presentation on the same computer you used to compose it. That means, you have to take the presentation file around.

Since you can insert memory intensive file formats to a POWERPOINT presentation it is not surprising that most presentations exceed a standard floppy disk in size. Another problem is that, if you want to take the presentation file – let’s say to a conference – you may not know whether there is POWERPOINT available.

Concerning these two problems POWERPOINT offers a smart solution: With PACK AND GO from the file menu you can save your presentation to multiple disks together with an POWERPOINT-viewer that allows you to perform the presentation even on computer systems without POWERPOINT available.

Figure 3-13: Pack and Go / Start



This wizard will guide you through the process without any problems. Just pay attention to include the viewer in step 5 if you are not sure whether you will find POWERPOINT available. After the wizard has collected all necessary information it will save the presentation to some floppies as executable files (*.exe).

On the system of your destination you simply insert the first disk of a set and start the *.exe file. This will produce an *.exe-file on the other computer that contains as well your presentation as the viewer. You can start the presentation by executing the *.exe-file.

4. MICROSOFT ACCESS

4.1. Databases: What they are and how do they work

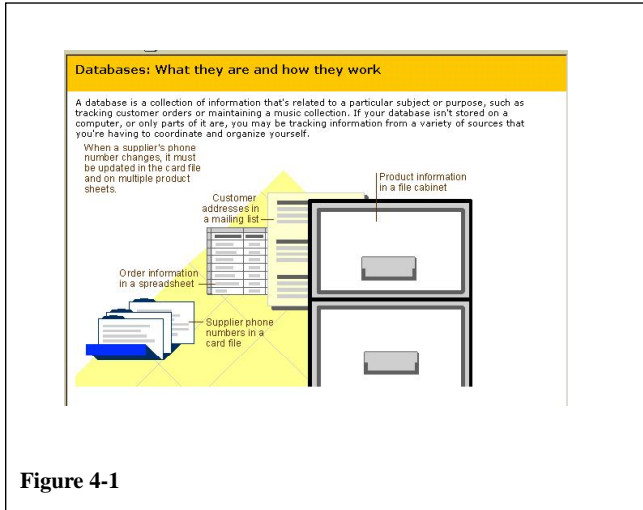


Figure 4-1

Databases: What they are and how they work

Using Microsoft Access, you can manage all your information from a single database file. Within the file, divide your data into separate storage containers called tables; view, add, and update table data by using online forms; find and retrieve just the data you want by using queries; and analyze or print data in a specific layout by using reports. Allow users to view, update, or analyze the database's data from the Internet or an intranet by creating data access pages.

Store data once in one table, but view it from multiple locations. When you update the data, it's automatically updated everywhere it appears.

Customers: Table

Customer ID	Company Name	City
BSBEV	B's Beverages	London
EASTC	Eastern Connection	London

Customers: Form

Customer ID: BSBEV
 Contact Name: Victoria Ashworth
 Company Name: B's Beverages

Orders by Customer: Report

Order ID	Total
10289	\$479.00

Figure 4-2

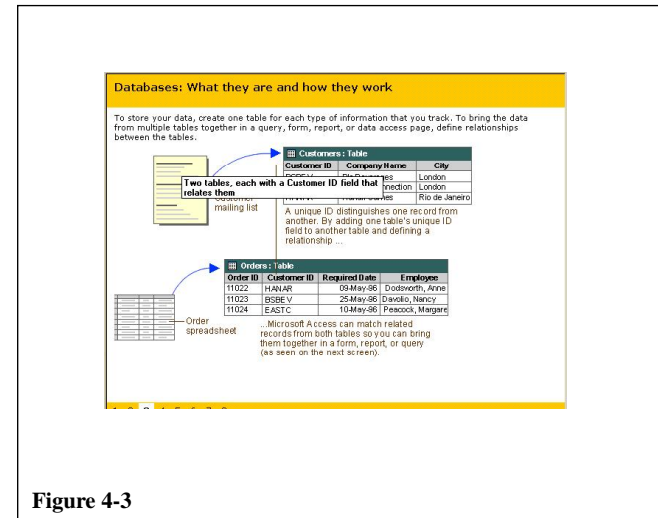


Figure 4-3

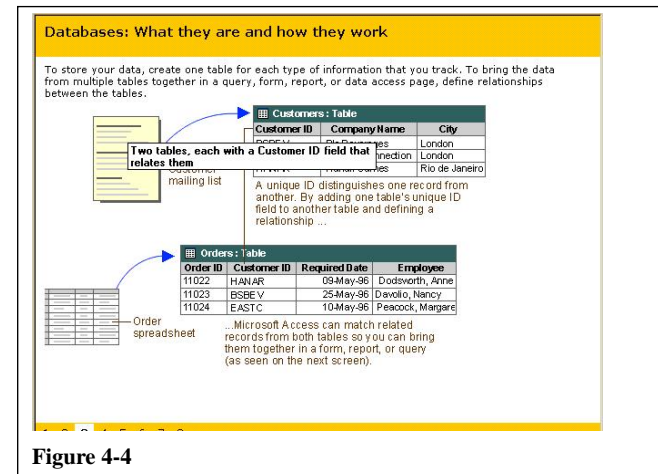
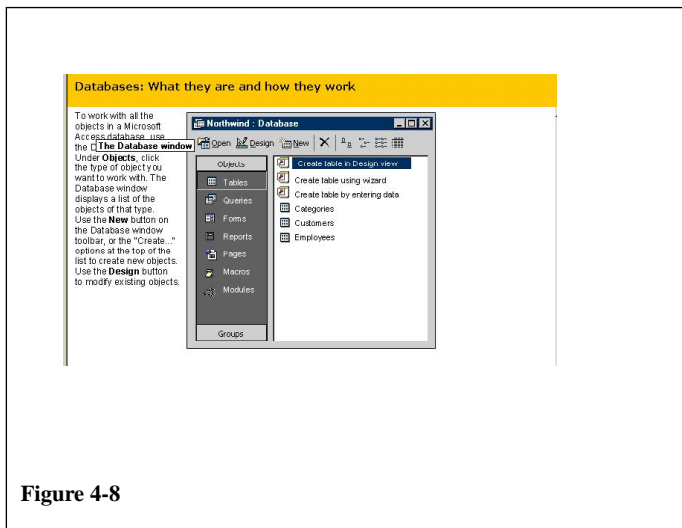
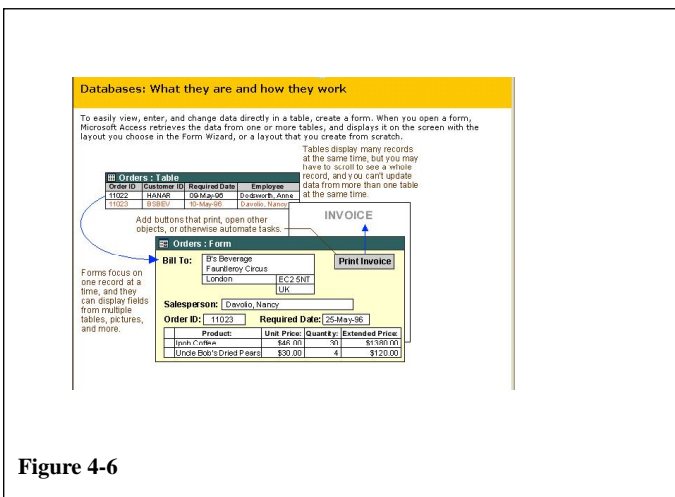
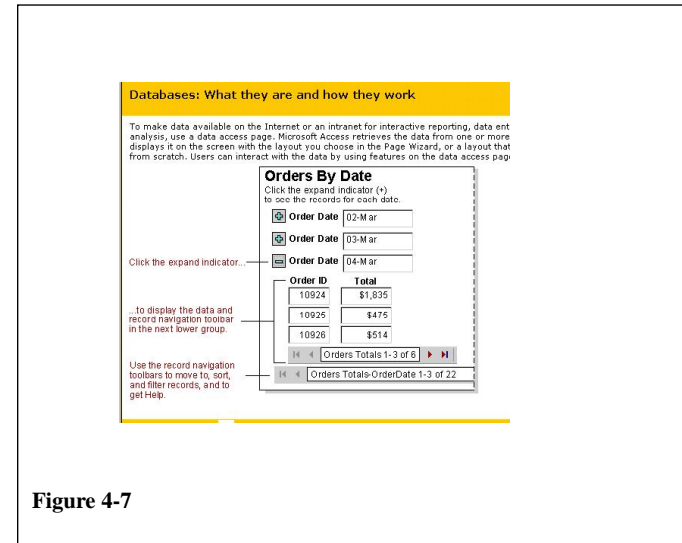
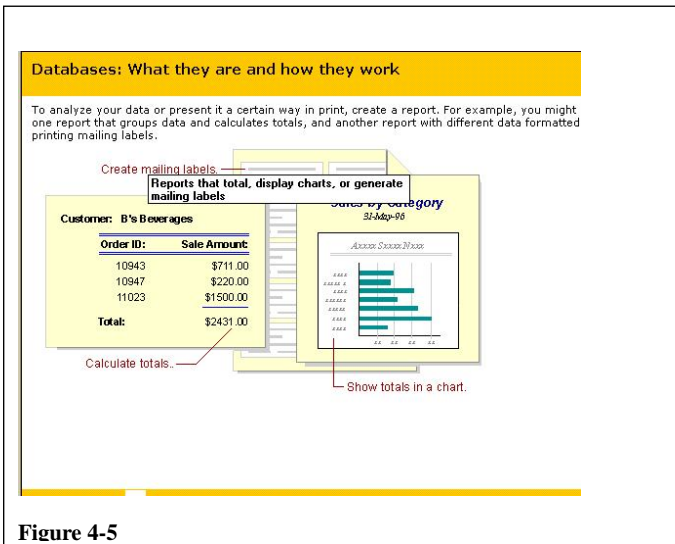


Figure 4-4



4.2. Tables: What they are and how do they work

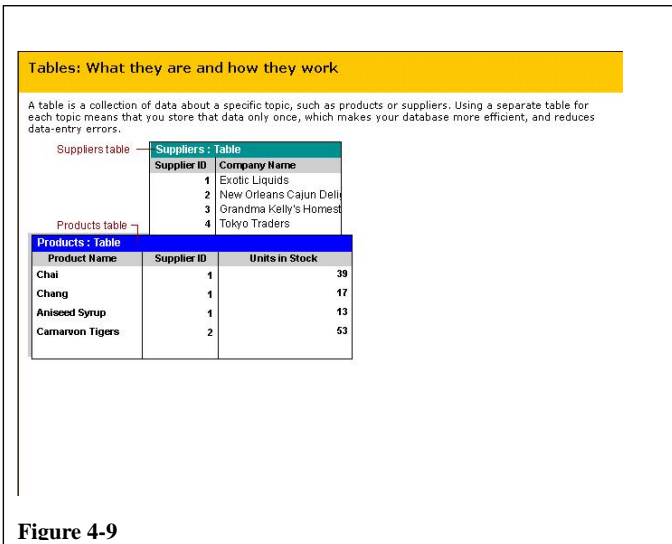


Figure 4-9

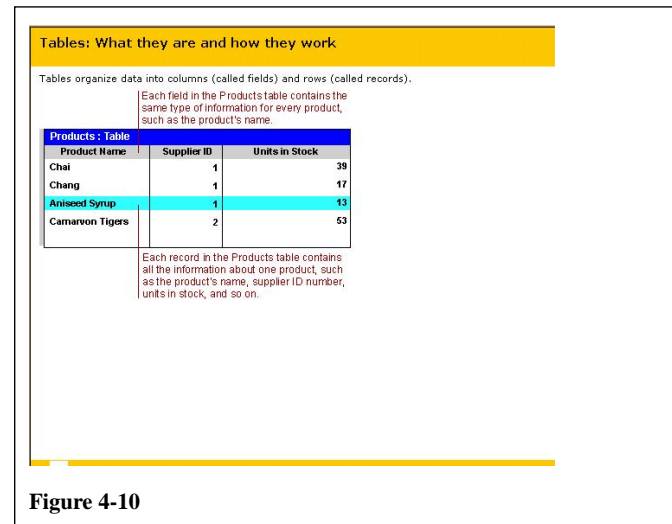


Figure 4-10

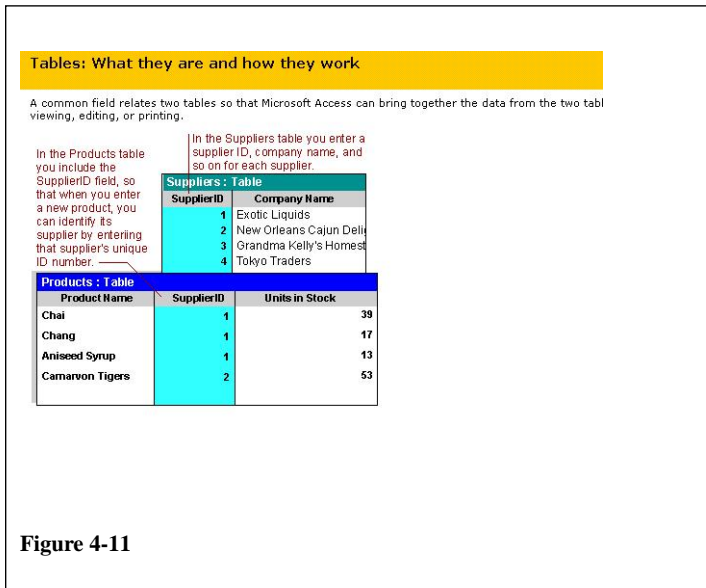


Figure 4-11

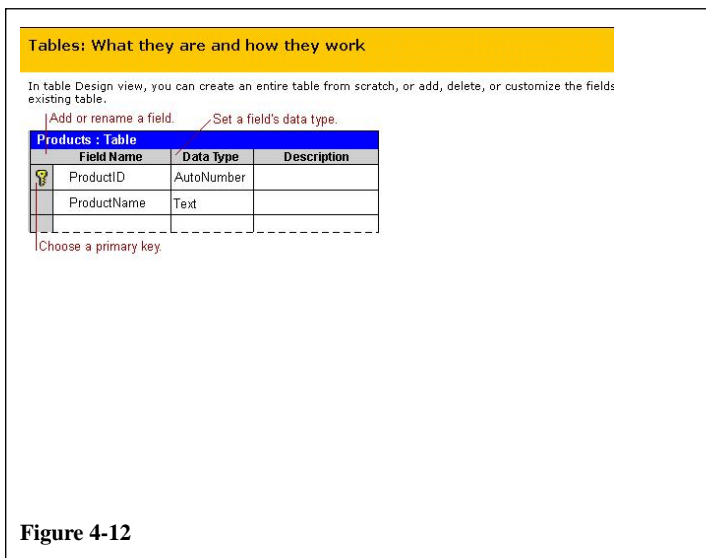
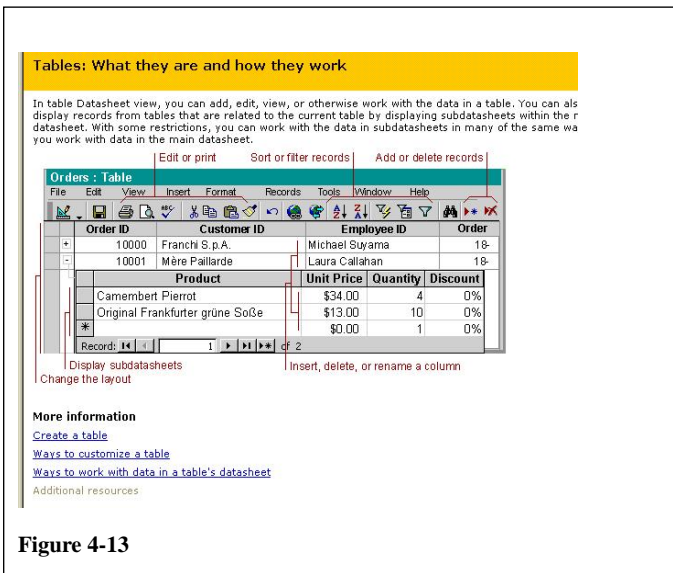
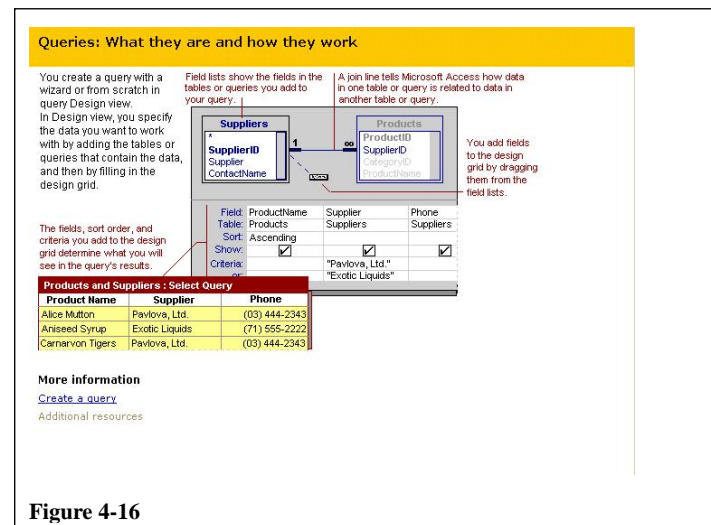
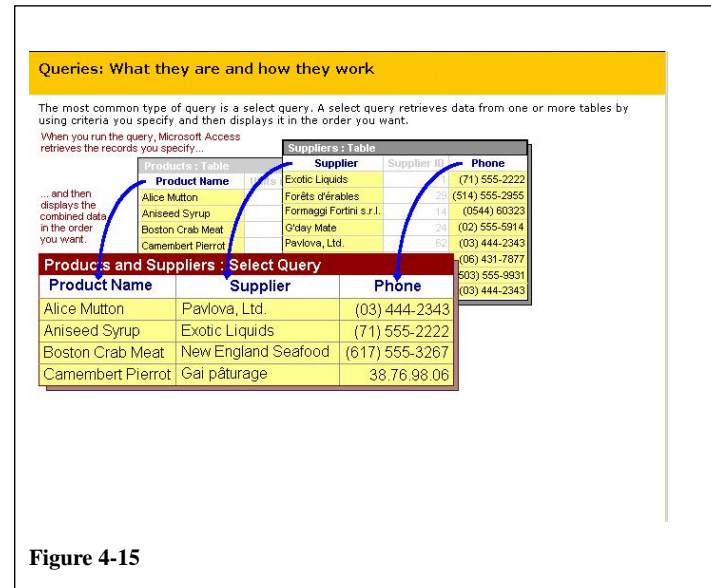
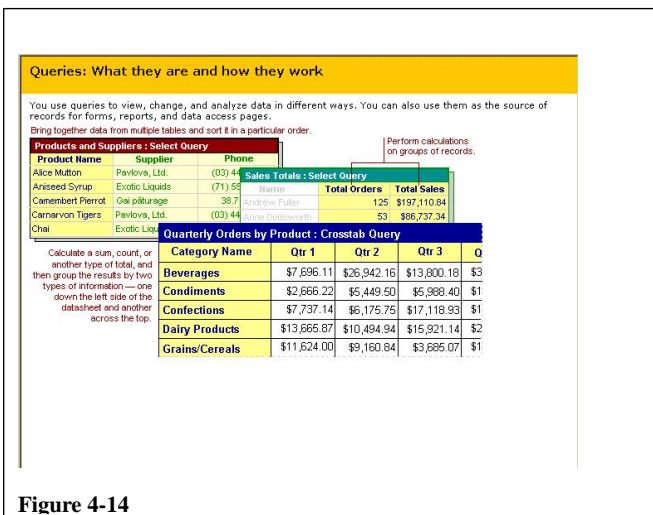


Figure 4-12



4.3. Queries: What they are and how do they work



4.4. Forms: What they are and how they work

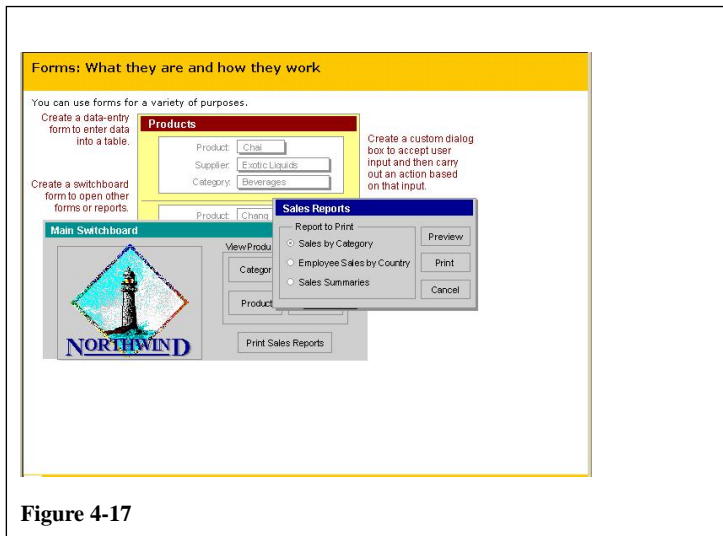


Figure 4-17

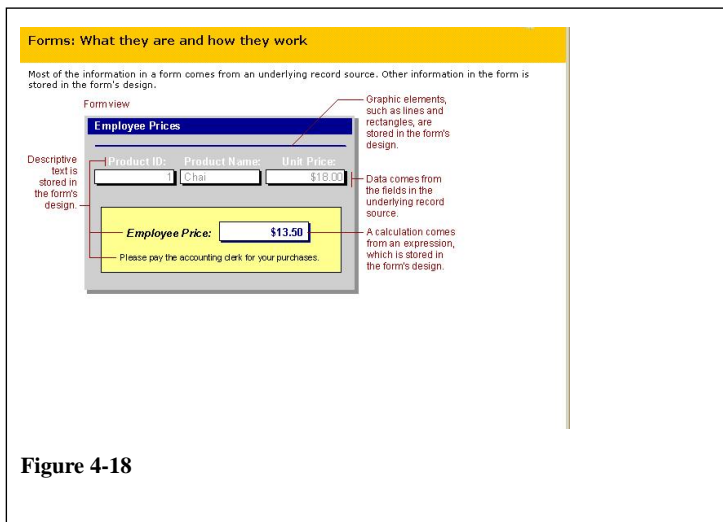


Figure 4-18

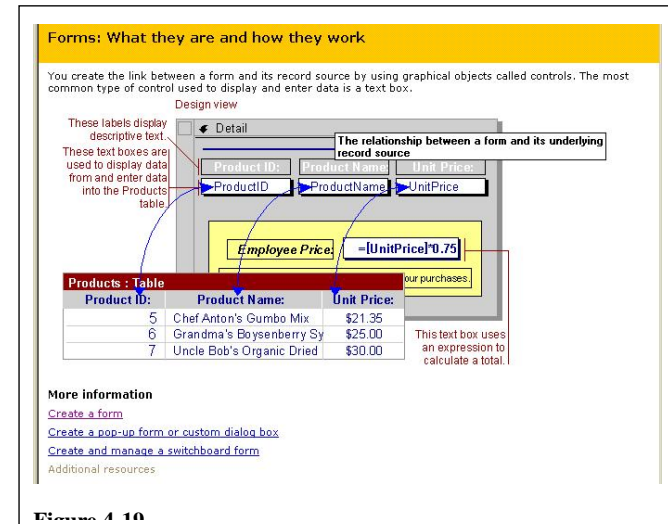


Figure 4-19

4.5. Reports: What they are and how they work

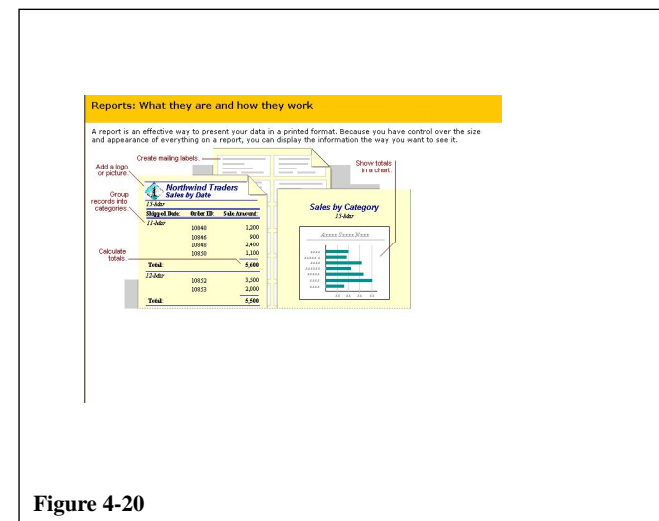


Figure 4-20

4.6. Data access pages: What they are and how they work

4.7. Macros: What they are and how they work

4.8. Modules: What they are and how they work

4.9. Relationships in a database

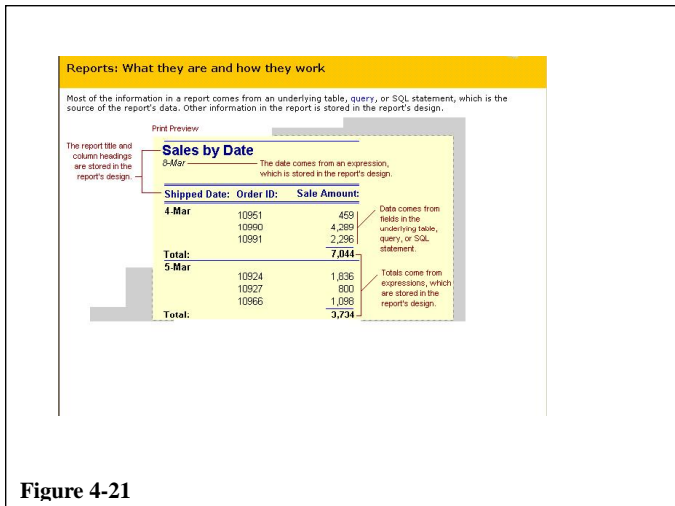


Figure 4-21

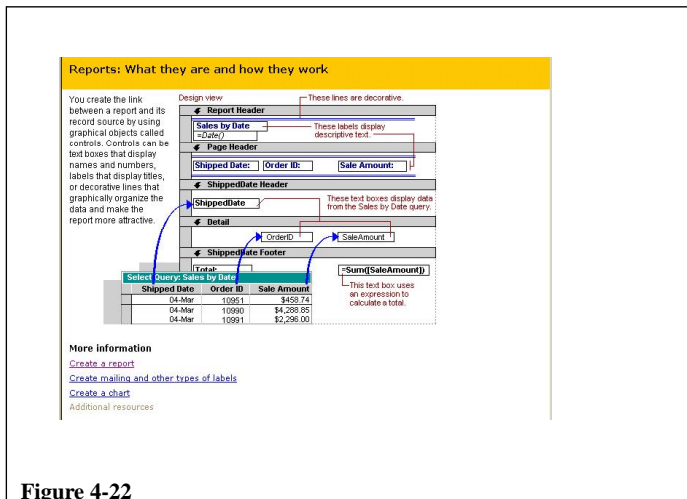


Figure 4-22

5. Hands on SPSS

5.1. A general view

The objective of this chapter is to give you a general view of the program. You should learn which are the main elements of SPSS, the types of windows and files SPSS uses

5.2. How to start and stop SPSS

- Click two times on the SPSS icon to start
- Choose File/Exit to leave the program

5.3. SPSS windows

SPSS uses the following most important types of window:

Data editor: A rectangular, spreadsheet-like display of the working data file. You can edit the data in the data window, add or delete variables, or change their attributes, except when the SPSS processor is modifying the data. Use File menu commands New and Open to clear data from the data editor, or to open an existing data file in the data editor. Use File menu commands Save or Save As when the data editor is active to save the data file. You cannot close the data editor, although you can minimize it. The file extension used by SPSS for this type of window is .sav.

Viewer. All statistical results, tables, and charts are displayed in the Viewer. You can edit the output and save it for later use. A Viewer window opens automatically the first time you run a procedure that generates output.

Results are displayed in the Viewer. You can use the Viewer to:

Browse results.

- Show or hide selected tables and charts.
- Change the display order of results by moving selected items.
- Move items between the Viewer and other applications.

The Viewer is divided into two panes:

The left pane of the Viewer contains an outline view of the contents.

The right pane contains statistical tables, charts, and text output.

You can use the scroll bars to browse the results, or you can click an item in the outline to go directly to the corresponding table or chart.

You can click and drag the right border of the outline pane to change the width of the outline pane.

Through the Viewer you can call various other editors

- **Pivot Table Editor.** Output displayed in pivot tables can be modified in many ways with the Pivot Table Editor. You can edit text, swap data in rows and columns, add color, create multidimensional tables, and selectively hide and show results.
- **Chart Editor.** You can modify high-resolution charts and plots in chart windows. You can change the colors, select different type fonts or sizes, switch the horizontal and vertical axes, rotate 3-D scatterplots, and even change the chart type.
- **Text Output Editor.** Text output not displayed in pivot tables can be modified with the Text Output Editor. You can edit the output and change font characteristics (type, style, color, size).

Syntax window with Syntax Editor: You can paste your dialog box choices into a syntax window, where your selections appear in the form of command syntax. You can then edit the command syntax to utilize special features of SPSS not available through dialog boxes. You can save these commands in a file for use in subsequent SPSS sessions.

If you have more than one open Viewer window, output is routed to the designated Viewer window. If you have more than one open Syntax Editor window, command syntax is pasted into the designated Syntax Editor window. The **designated** windows are indicated by an exclamation point (!) in the status bar. You can change the designated windows at any time. The designated window should not be confused with the **active** window, which is the currently selected window. If you have overlapping windows, the active window appears in the foreground. If you open a new Syntax Editor or Viewer window, that window automatically becomes the active window and the designated window.

5.4. Example: The questionnaire of the Munich Centre's sample survey

To have a very simple data set to show the effect and handling of the SPSS commands, please consider to following very simple (if not simplistic) questionnaire: The sample survey of the Training Course of the Munich Centre:

The questionnaire

We would like to know more about your general information about Germany in your country. Please answer the following questions

1. Which is your sex?	male	
-----------------------	------	--

Fill in your personal copy of the questionnaire, then start SPSS to start working with these simple data ✓

2.4. The data file

Open a new data file:
File/New/Data

Then a new window is opened with the following name: *Untitled*

For the five variable we will define the following names:

Question	Name of variable
No of questionnaire	nrquest
1.	sex
2.	age
3.	per_germ
4.	com_germ

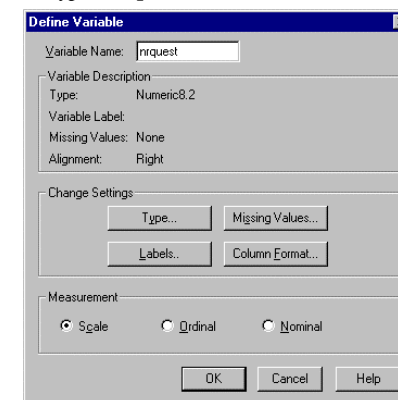
This is your first interactive exercise in SPSS. This should be done together with and guided by your trainer. Try to follow these steps exactly as indicated. You can do the stand-alone exercises later on in your own way and pace.

Start SPSS, Select [Type in data] if the initial help-screen still appears, click anywhere on the first column of the window called *Untitled*

Choose [Data/Define Variable]

1. Then a dialog box will appear and you can indicate the name of the variable.

Type: "**nrquest**" into the indicated edit-box.



2. Fill then the names of the other variables. Enter the values of the observations (each questionnaire represents one observation). Use the keyboard or the mouse to highlight one cell, then type the content.. [ENTER] accepts the value and moves the cursor one line downwards. [TAB] accepts the value and moves the cursor one column to the right.

Figure 15: Define Variables

We would like to know more about your general information about Germany in your country. Please answer the following questions

	female	<input type="checkbox"/>
2. Which is your age?		<input type="checkbox"/>
3. Who is the contemporary German personality best known in your country (Indicate only one person) ?	Gerhard Schröder	<input type="checkbox"/>
	Boris Becker	<input type="checkbox"/>
	Claudia Schiffer	<input type="checkbox"/>
	Franz Beckenbauer	<input type="checkbox"/>
	Helmut Kohl	<input type="checkbox"/>
	Steffi Graf	<input type="checkbox"/>
	None/Other	<input type="checkbox"/>
	Don't know	<input type="checkbox"/>
4. Which is the German company enjoying the best reputation in your country (Indicate only one company) ?	Mercedes-Benz	<input type="checkbox"/>
	Carl-Duisberg Gesellschaft	<input type="checkbox"/>
	Bayer	<input type="checkbox"/>
	MAN	<input type="checkbox"/>
	BMW	<input type="checkbox"/>
	Siemens	<input type="checkbox"/>
	None/Other	<input type="checkbox"/>
	Don't know	<input type="checkbox"/>

Even if this is not a very realistic example, it should help to understand how SPSS works, how data are entered and subsequently analyzed.

Enter your personal data. Choose an ordinal numbering for the variable *nrquest* (start from 1 for the first questionnaire, 2 for the second etc....). Code the other data according to their position in the questionnaire (first answer is 1, second is 2 etc. For example: a male respondent who answers, that Claudia Schiffer is the best known German personality and Siemens enjoys the best reputation in your country codes the replies to the questions 1, 3 and 4 as 1, 3 and 6 respectively. The age is coded directly in number of years.

Pass your questionnaire to your neighbor on your right hand and fill in the questionnaire of your left hand neighbor until all questionnaires are entered to your data file

2.5. Save the data file

Save your work:
Choose [File/Save As]
then type the file name, e.g.: *quest01*. The file will be saved as *quest01.sav*. The table and window will appear like this.
Watch where you save your work. It is a good idea to create a separate directory for your data file. You should do this using the WINDOWS-Explorer.

	nrquest	sex	age	per_germ	com_germ
1	1.00	2.00	18.00	1.00	4.00
2	2.00	2.00	41.00	2.00	4.00
3	3.00	1.00	53.00	5.00	1.00
4	4.00	2.00	26.00	4.00	2.00
5	5.00	1.00	27.00	5.00	6.00
6	6.00	1.00	40.00	6.00	2.00
7	7.00	1.00	66.00	3.00	6.00
8	8.00	2.00	26.00	4.00	6.00
9	9.00	1.00	27.00	5.00	6.00
10	10.00	1.00	34.00	6.00	2.00

Figure 16: Saving your work

Generally SPSS will save your work in the actual directory, e.g. the SPSS directory where all the system files are stored. This is no good place for your data files or later for your results.

2.6. Listing cases

You want to see, what has been stored in your data file .

Listing the captured observations:

Choose [Statistics/Summarize/ Case Summaries...]. Select one or more variables. This is done selecting from the following dialog box

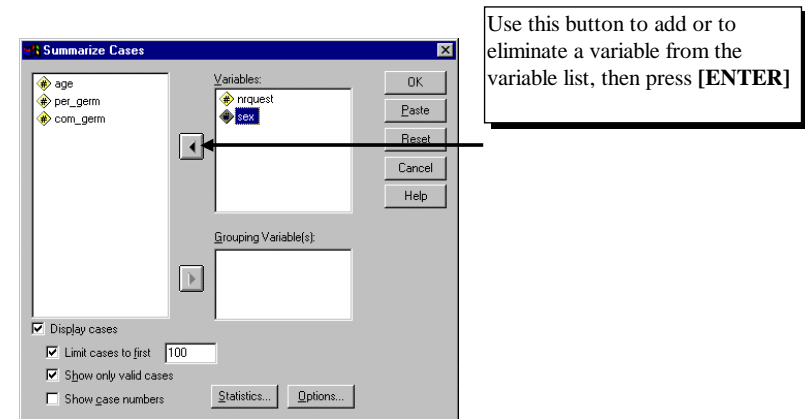


Figure 17: Selecting variables for listing

The result of this operation is listed in SPSS Viewer. You can select the complete output in the right-hand window or select individual items of the output, like the Case Summaries only .

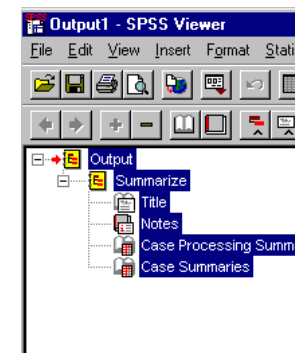
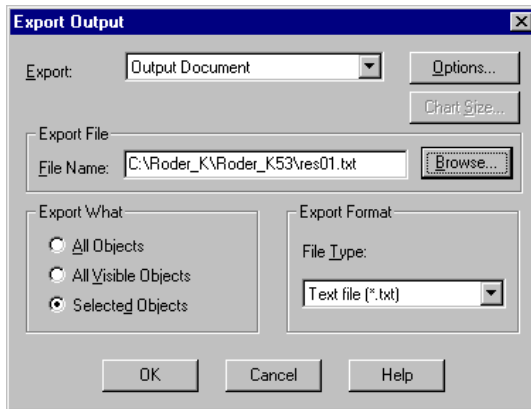


Figure 18: The SPSS Viewer

You can save this output as a Viewer document (.spo) which will be accessible from SPSS only. If you want to work with the results in other programs, e.g. a word processor for compiling a report, you can save the results directly by selecting the appropriate tables via the Clipboard.

You can also save this file like a text file, e.g. like res01.txt. You do this by choosing [File/Export] and filling the corresponding dialog box:



The content of this file is the following:

Case Summaries

Figure 19: Exporting results

	NRQUEST	SEX	AGE	PER_GERM	COM_GERM
1	1.00	2.00	18.00	1.00	4.00
2	2.00	2.00	41.00	2.00	4.00
3	3.00	1.00	53.00	5.00	1.00
4	4.00	2.00	26.00	4.00	2.00
5	5.00	1.00	27.00	5.00	6.00
6	6.00	1.00	40.00	6.00	2.00
7	7.00	1.00	66.00	3.00	6.00
8	8.00	2.00	26.00	4.00	6.00
9	9.00	1.00	27.00	5.00	6.00
10	10.00	1.00	34.00	6.00	2.00
11	11.00	2.00	33.00	1.00	2.00
12	12.00	1.00	45.00	2.00	1.00
13	13.00	2.00	88.00	1.00	4.00
14	14.00	1.00	54.00	5.00	1.00
15	15.00	2.00	14.00	3.00	2.00
16	16.00	2.00	29.00	6.00	5.00
17	17.00	1.00	44.00	4.00	6.00
18	18.00	1.00	22.00	5.00	6.00
19	19.00	1.00	9.00	6.00	2.00
20	20.00	1.00	23.00	1.00	1.00
21	21.00	1.00	39.00	6.00	5.00
22	22.00	1.00	61.00	3.00	6.00
23	23.00	2.00	36.00	4.00	6.00
24	24.00	2.00	36.00	4.00	6.00
25	25.00	2.00	13.00	2.00	2.00
26	26.00	2.00	78.00	1.00	4.00

27	27.00	1.00	64.00	5.00	1.00
28	28.00	2.00	15.00	4.00	2.00
29	29.00	2.00	33.00	1.00	2.00
30	30.00	1.00	47.00	2.00	1.00
Total	N	30	30	30	30

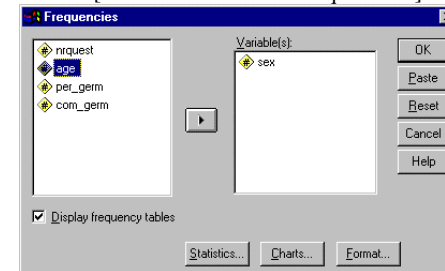
a Limited to first 100 cases.

5.5. The analysis of frequency



After this, start with the first simple statistical analysis: How many women an men are in the group of respondents. Analyze the frequency of sex:

Choose [Statistics/Summarize/Frequencies..].



Then choose the variables using the dialog box and choose the option [Charts]: With the choice of Bar Chart(s) and [Continue] in the second and the choice of [Paste] in the first window you get the following results: a syntax file, which you can save (syn01.sps) and another result: tables and graphics. The command to analyze the frequency will be executed by choosing [Run/All] or pressing the button in the syntax window,

Figure 20: Choosing variables for frequencies

if the cursor is positioned somewhere on the SPSS command line and all of the syntax file has been selected with your mouse.

The content of the syntax file:

```
FREQUENCIES
  VARIABLES=sex
  /BARCHART FREQ
  /ORDER ANALYSIS .
```

The content of the output file:

The two tables:

Statistics

```
SEX
      N      Valid 30
      Missing  0
```

```
SEX
      Frequency Percent      Valid      Cumulative
      Valid      1.00      16      53.3      53.3      53.3
```

2.00	14	46.7	46.7	100.0
Total	30	100.0	100.0	

The content of the graphics file

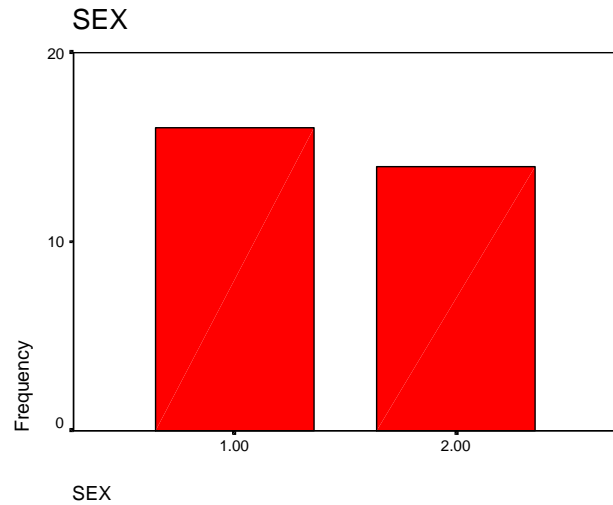


Figure 21: Resulting graph of FREQUENCIES command

High-resolution graphics allows interpretation of results with more ease. Especially statistical publication and on the spot information for a less professional public asks for extensive use of graphical representation of statistical data.

Save the output file in the SPSS Viewer

1. in the SPSS format (.spo file) *or*
2. export in Internet file format (.htm). The text files will be in .htm format, the graph in JPEG (.jpg) format *or*
3. export in Text file format (.txt). The text part will be saved in ASCII .txt format, the graph will be given a new file name, internally assigned in .jpg format *or*

4. Save the content of the Viewer and the output file through the Clipboard to store it in another program environment.

Exercise: The Cameroon Household Survey:

In the annex you will find a description of the Cameroon Household Survey. The data file you will find under the name of CHS96_01.SAV. This file contains raw data of a sample of members of households of section 1 and section 10.


Use SPSS to handle and analyze the data:

1. Start the read the data file. Count the number of variables and observations
2. List the content of 50 observations
3. Print the frequency count of men and women in the universe.
4. Rename variables, if necessary or appropriate (e.g. s1q3 = sex)
5. Find out which variable contains the information about sex and print a bar graph of the variable.
6. Print a histogram of the variable *age* (this variable might not yet exist, see exercise 4.) and comment the differences of the distribution of this variable against a normal distribution.

Additional exercise:

7. Create a new variable BRID, which identifies exactly the household within the sample.

5.6. The transformation of variables

 Before being able to analyze the data, usually some preparatory work is necessary, the transformation of variables and their respective values. You should know the possibilities and the reasons why these transformations are necessary and how they are done.

5.7. Defining variables

To define new variables or change the format of existing variables, double-click a column in the Data Editor, or select a cell in the column and choose [Define Variable] from the [Data] menu. You can also avoid to define variables and type data directly into the specified cells. Then the variable names will be provided by SPSS by default: the first variable named will be VAR00001, the second VAR00002 etc.

5.8. Variable names in SPSS

SPSS variable names can contain up to 8 characters. The first character of a variable name that you define must be a letter or @. Each subsequent character must be a letter, a numeral, a period, an underscore, or one of \$#@, except that the final character cannot be a period. Uppercase and lowercase letters are equivalent in variable names. You cannot use as a variable name: ALL, AND, BY, EQ, GE, GT, LE, LT, NE, NOT, OR, TO, WITH.

5.9. Variable Types

The following are valid variable types:

Numeric
Comma
Dot
Scientific Notation
Date
Dollar
Custom Currency
String

The two types most frequently used are described briefly below:

Numeric Variable

Defines a numeric variable whose values are displayed in standard numeric format, using the decimal delimiter specified in the International control panel. The data editor accepts numeric values in standard format; or in scientific notation, provided that both the E and the sign of the exponent are entered (1E+1 but not 1E1).

String Variable

Defines a string variable, whose values can contain any characters up to the defined length. Upper- and lower-case letters are considered distinct. Some SPSS commands cannot use string variables, since calculations with strings are not defined. Strings defined as longer than 8 characters in length ("long strings") are restricted in use.

There are functions to convert numeric into string variables and vice versa. Even though you will learn about SPSS functions later, these two important ones are explained below:

STRING(numexpr,format) String Returns the string that results when numexpr is converted to a string according to format. The second argument format must be a format for writing a numeric value.

NUMBER(strexp,format) Numeric Returns the value of the string expression strexp as a number. The second argument, format, is the numeric format used to read strexp. Thus if name is an 8-character string containing the character representation of a number, **NUMBER(name, f8)** is the numeric representation of that number.

5.10. Missing values

There are two types of missing values in SPSS, user-missing values and system-missing values.

Any blank numeric cells in the data rectangle are assigned the system-missing value, which is displayed as a period. You can also assign values that identify information missing for specific reasons, and then instruct SPSS to flag these values as missing. System-missing and user-missing values are handled specially by SPSS statistical procedures and by data-transformation commands.

5.11. Variable labels

A variable label contains descriptive text of up to 120 characters used to label SPSS output. The variable label typically appears beside the variable name, and serves to describe the meaning or origin of the variable. (The output from many procedures displays only about 40 characters from the variable label.)

Choose: [Data/Define Variable]

This command is available when the Data Editor is the active window. You can either define a new variable, or change the format of an existing variable. Type a variable name for the variable. (To rename an existing variable, replace the existing name.) SPSS uses some special variable names, and you should avoid these.

To change the type or display format, select [**Type**].
To specify missing values, select Missing [**Values**].

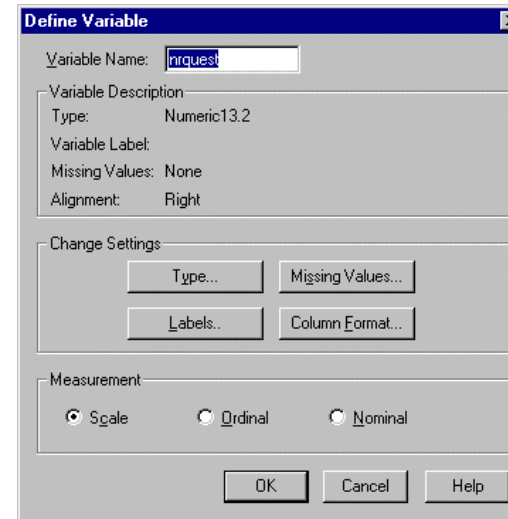


Figure 22: Defining Variables

To supply a variable label or value labels, select [**Labels**].
To change the alignment or width of the variable's column in the Data Editor, select [**Column Format**].

Now you will add labels to your data values. Please use the ones in the following syntax file

From now on, all SPSS commands will be executed as if a syntax file had been edited and run.

You will perform these tasks usually selecting the menu choices and paste your choices into the Syntax Editor. You may continue

to do so in further exercises. As you go along in this course, the SPSS syntax should become more familiar and 'writing' a SPSS program will be the most natural thing to you. At least, this is what the author and probably trainer hopes so.

Example: The MC sample survey

We will apply the first transformation to the initial data file. Subsequently all syntax files will have a header like below. For the time being you will enter the labels using the menu commands. Later you will learn to use syntax files.

```
*****
* Syntax file           : ex03_1.sps
* Data File charged    : QUEST01.SAV
* Data File saved      : QUEST02.SAV
* Output File          : -----
```



```
* Graphics File      : -----
* Comments          :
* Comments          :
* Author            : CDG-FZ München/klaus röder
*****
GET FILE='C:\Roder_K\Roder_K53\QUEST01.SAV'.
VARIABLE LABELS
nrquest 'Number of Questionnaire'
/sex 'Sex'
/age 'Age'
/per_germ 'Best known German Personality'
/com_germ 'Best known German Company'
.
```

```
VALUE LABELS
/sex 1 'male' 2 'female'
/per_germ 1 'Gerhard Schröder'
         2 'Boris Becker'
         3 'Claudia Schiffer'
         4 'Franz Beckenbauer'
         5 'Helmut Kohl'
         6 'Steffi Graf'
         7 'None/Other'
         8 "Don't know"
/com_germ 1 'Mercedes-Benz'
         2 'Carl-Duisberg Gesellschaft'
         3 'Bayer'
         4 'MAN'
         5 'BMW'
         6 'Siemens'
         7 'None/Other'
         8 "Don't know"
.
```

```
SAVE OUTFILE='C:\Roder_K\Roder_K53\QUEST02.SAV'.
```

The path in *gray* may vary in your exercise. Please change to your path names so that this example executes on your computer and your path of data files allocated for this course.

5.12. Compute variables

Menu: [Transform/Compute]

Command Language: *COMPUTE, IF*

To compute values for all cases, type the name of a single Target Variable. It can be an existing variable, or a new variable to be added to your working data file. Build an Expression for the values that should be assigned to that variable.

To compute values selectively, for cases satisfying some logical condition, select the **[If]** pushbutton as well. If the Target Variable is or will be a string variable you must select **[Type] & [Label]** to indicate its length. If you are computing a new variable, you can select **[Type] & [Label]** to specify a variable label for it.

To build an expression, either paste components into the *Expression* field, or place the cursor there and type.

Paste variable names by copying them from the source list at the left.

Paste numbers and operators from the calculator pad.

Paste functions by copying them from the function list at the right.

String constants must be enclosed in quotation marks or apostrophes.

Numeric constants must be typed in American format, with the dot as a decimal delimiter.

If the Transformation Options setting (in Edit/Preferences) is *Calculate values immediately* (the initial setting), SPSS reads the data file and carries out the data transformation when you select **[OK]**. If the setting is *Calculate values before used*, SPSS carries out the transformation as the next procedure reads the data, so the Data Editor will not reflect the transformation until then. Additional features are available in the SPSS command language.

5.13. Recode variables

Recode into Same Variables

Menu: [Transform/ Recode/into Same Variables...]

Command language: *RECODE*

Move the variables that you want recoded according to the same specifications into the Variables list box. Select Old and New Values and specify how to recode the variable(s). If you want to recode only the cases satisfying a logical condition, select **[If]** and specify the condition.

Recode into Different Variables

Menu: [Transform/ Recode/into Different Variables...]

Command Language: *RECODE...INTO*

Move the variables of which you want copies recoded according to the same specifications into the *Input Variable -> Output Variable* list box. Select each variable individually in that box and, in the Output Variable group, type the Name (and optional Label) of a new Output

Variable. Select [**Change**] to apply the name and label. If you modify the name or label, you must select [**Change**] again.

Select [**Old and New Values**] and specify how to recode the variable(s). If you want to recode only the cases satisfying a logical condition, select [**If**] and specify the condition.

See end of paragraph 3.6 for Transformation Options.

5.14. Transpose variables

Menu:[Data/Transpose]
Command Language: *FLIP*

Transpose transposes the rows and columns in the data file so that cases become variables and variables become cases. Select at least one variable. These variables are columns in the working data file, and will be transposed into cases (rows) in the new data file. Variables that you do not select will be dropped. SPSS reports the variable names in the transposed file, and creates a variable *CASE_LBL* containing the old variable names.

If the working data file contains an ID or name variable with unique values, you can move it into the *Name Variable* box, and its values will be converted into variable names for the transposed data file

5.15. Sort cases

Menu:[Data/Sort Cases]
Command Language: *SORT CASES*

Select a variable and move it into the *Sort by* box. To sort in descending order, select the variable in the *Sort by* box and press [**Descending**]. You can select additional sort variables. Each additional sort variable specifies the order of cases within the group that have identical values for all preceding sort variables. Ascending order is represented by (*A*) in the list of sort variables. String variables are sorted alphabetically. Descending order is represented by (*D*) in the list of sort variables. String variables are sorted in reverse alphabetical order.

The dialog boxes for the Data/Split File menu and the Statistics/Summarize/ Split File menu can automatically sort your data file, so it is not necessary to do so yourself before those commands.

5.16. Select cases

Menu: [Data/Select Cases]
Command Language: *SELECT IF; SAMPLE; FILTER; USE*

Choose one of the alternatives in the Select group:

All cases indicates that you wish to process all of the cases in your working data file.

If condition is satisfied indicates that you wish to process only the cases satisfying a logical condition. If you have specified a condition, it appears beside the [**If**] Pushbutton. If not, or if you want to change the condition, click on [**If**]. This alternative sets the variable *filter_\$*, discussed below.

Random sample of cases indicates that you wish to select cases randomly for processing. If you have entered sampling specifications they appear beside the [**Sample**] Pushbutton. If not, or if you want to change them, click on [**Sample**]. This alternative sets the variable *filter_\$*, discussed below.

Based on time or case range indicates that you wish to process only those cases falling within a range of specified dates. Click on [**Range**] to define the dates to be included in processing.

Use filter variable indicates that you want to use the values of an existing numeric variable to control case filtering. Select the variable from the list at the left. Cases for which the filter variable has the value 0 are excluded from analysis.

Choose one of the alternatives in the *Unselected cases are* group:

Filtered excludes unselected cases from processing, but retains them in the data file. You can re-enter this dialog box and change the Select alternative to restore the cases.

Deleted drops the cases from the data file. To recover them, you must Open the original data file again (assuming you have not replaced it with the reduced version).

If you execute this command with the [**If**] or [**Sample**] alternative and request that unselected cases be filtered, a variable named *filter_\$* is calculated as 1 (selected) or 0 (not selected). You can edit the values of *filter_\$* in the Data Editor; cases with a value of 0 for a filter variable are excluded from analysis, while cases with any other valid numeric value are included. However, be aware that this variable holds the current filtering status. If you do choose to edit its value, this affects which cases are filtered.

✓ In the following example we will apply some of the mentioned commands:

1. Create an age groups in steps of 10 (e.g. age group value 1 for all ages less than or equal to 10)
2. Attach labels to these values (e.g. ' ≤ 10 ') for group value 1
3. Name the missing values and attach labels
4. Sort cases by age groups
5. Apply a filter (all cases aged under or equal to 30)

Example: The MC sample survey

Let us apply some more transformation commands on the initial data set of the MC sample survey. The commands are given in a syntax file, for the beginning the menu options should be used, but the [**Paste**] button can be used to see the output in SPSS syntax.

```
*****
* Syntax file       : ex03_2.sps
* Data File charged : QUEST02.SAV
* Data File saved  : QUEST3.SAV
* Data File saved  : QUEST4.SAV
* Output File     : -----
* Graphics File   : -----
* Comments        :
* Comments        :
* Author          : CDG-FZ München/klaus röder
*****
GET FILE='C:\RODER_K\RODER_K53\QUEST02.SAV'.
COMPUTE gr_age = age .
VARIABLE LABELS gr_age "Age group" .
IF (age <= 10) gr_age = 1 .
IF (age > 10 & age <= 20) gr_age = 2 .
IF (age > 20 & age <= 30) gr_age = 3 .
IF (age > 30 & age <= 40) gr_age = 4 .
IF (age > 40 & age <= 50) gr_age = 5 .
IF (age > 50 & age <= 60) gr_age = 6 .
IF (age > 60 & age <= 70) gr_age = 7 .
IF (age > 70 ) gr_age = 8 .
MISSING VALUES
/gr_age (-1 -2).
VALUE LABELS
/gr_age
-1 'No response'
-2 'Age unknown'
1 ' <= 10'
2 '10< <=20'
3 '20< <=30'
4 '30< <=40'
5 '40< <=50'
6 '50< <=60'
7 '60< <=70'
8 '70< '
.
EXECUTE .
SAVE OUTFILE='C:\RODER_K\RODER_K53\QUEST03.SAV'.
SORT CASES BY gr_age .
SPLIT FILE
  BY gr_age .

SORT CASES BY
  age (A) .

USE ALL.
COMPUTE filter_$=(gr_age <= 3).
VARIABLE LABEL filter_$ 'gr_age <= 3 (FILTER)'.
```

```
VALUE LABELS filter_$ 0 'Not selected' 1 'Selected'.
FORMAT filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE .
FILTER OFF.
USE ALL.
SELECT IF(gr_age <= 3).
EXECUTE .
SAVE OUTFILE='C:\RODER_K\RODER_K53\QUEST04.SAV'.
```


Exercise: The Cameroon Household Survey:

In the annex you will find a description of the Cameroon Household Survey. Use the data file you used before in the exercise of chapter 1 or use CHS96_02.SAV

Use SPSS to transform the data:

1. Add variable and value labels for some selected variables (*sex*, *relation to head of household* etc.).
2. Define missing values for at least one variable (candidates are *age* and *Why not attending School*).
3. Define a new variable *gr_age*
4. Compute values for *gr_age* for each group comprising 10 years (*gr_age*= 1 for *age* between 0 and 10 years, *gr_age* = 2 for *age* between 10 and 20 years etc).
5. Recode values for *sex* (1=0 and 2=1).
6. Select only respondents older than 20 years and count the frequency of *Attending School*.

5.17. The transformation of data files

 Before data analysis can start, very often data files have to be prepared and transformed. It may be that files have to be imported from different sources, e.g. data bases like ORACLE, EXCEL or xBASE, sometimes several files have to be added, sometimes they have to be merged. In this chapter you should learn how this transformation is done and why this might be necessary.

5.18. Merging Files: General considerations

The dialog-box interface to the Merge Files facility combines two data files: the working data file, and a data file saved in SPSS format (called the "external file"). Two types of merge are supported:

1. Add Cases adds all of the cases from the external file to the end of the working data file. The new file includes whatever variables were shared by the two files, by default. You can choose to include variables that were in only one file; cases coming from the other file will have system-missing or blank values for such variables. You can also exclude variables.
2. Add Variables adds some or all of the variables in the external file to corresponding cases in the working data file. You can define "corresponding" cases in any of three ways:

- a. Case order. If you do not specify a key variable, cases are simply aligned according to their order in both files. This is risky unless you are certain that case orderings in the two files correspond, so usually type b. or c. of merging is selected.
- b. Key variable(s). If you specify a key variable, cases are aligned according to the values of that variable, which typically has unique identifying values. If necessary, you can use more than one key variable to specify the correspondence of cases. Both files must be sorted by the key variable(s). Unmatched cases in either file enter the new working data file, with system-missing or blank values for the variables that come from the other file.
- c. Table lookup. If you specify a key variable and indicate that one of the files is a "table," cases are aligned according to the key variable, but cases appearing only in the table do not contribute to the new working data file. The table simply provides data for new variables added to the other file. Both files must be sorted by the key variable(s).

If you have selected [*Calculate values before used*] in [Edit/Options.../Data], the files are not merged until you execute another command, such as a statistical procedure.

5.19. Merge Files: Add Cases

Menu:

[Data/Merge Files/Add Cases]

Command Language: *ADD FILES*

An initial dialog box asks you to identify the file from which you want to add cases.

The list box at the right, *Variables in New Working Data File*, contains all the variables shared by both files. If this list is satisfactory, press **[OK]**. The list box at the left, *Unpaired Variables*, contains variables that could not be matched across the files, because they had different names or different types. Variables marked with (*) are in the working data file. Variables marked with (+) are in the external file. Select one of the following for instructions:

Including a variable that is in one file only.

Including a variable with two different names.

Splitting a variable into two variables (giving different names for variables with identical names from two sources).

Excluding a variable.

Renaming a variable in the left list box.

Renaming a variable in the right list box.

	nrrequest	sex	age	per_germ	com_germ	g
1	4.00	1.00	18.00	.00	3.00	
2	14.00	1.00	41.00	1.00	3.00	
3	23.00	.00	53.00	4.00	.00	
4						
5						
6						
7						
8						
9						
10						

Figure 23: File QUE04_1.SAV

If you select *Indicate case source as variable* a variable is added to the new working data file which has the value 0 for cases from the original working data file, and the value 1 for cases from the external file.

	nrrequest	sex	age	per_germ	com_germ	g
1	4.00	.00	40.00	5.00	Best Known German Campaign	
2	14.00	.00	66.00	2.00	5.00	
3	23.00	.00	27.00	4.00	5.00	
4	40.00	.00	34.00	5.00	1.00	
5	41.00	1.00	33.00	.00	1.00	
6						
7						
8						
9						
10						

Figure 24: File QUE04_2.SAV

	nrrequest	sex	age	per_germ	com_germ	g
1	4.00	1.00	18.00	.00	3.00	
2	14.00	1.00	41.00	1.00	3.00	
3	23.00	.00	53.00	4.00	.00	
4	4.00	.00	40.00	5.00	1.00	
5	14.00	.00	66.00	2.00	5.00	
6	23.00	.00	27.00	4.00	5.00	
7	40.00	.00	34.00	5.00	1.00	
8	41.00	1.00	33.00	.00	1.00	
9						
10						

Figure 25: Result of ADD FILES.. File QUE04_03

5.20. Merge Files: Add Variables

Menu: [Data/Merge Files/Add Variables]
Command Language: *MATCH FILES*

An initial dialog box asks you to identify the file from which you want to add variables.

Specify the type of match:

If you want to match the cases in the two files by sequential order alone, leave *Match cases on key variables* in sorted files unchecked.

If both files are sorted by one or more "key variables" and you want to match the cases according to their values: select *Match cases on key variables* in sorted files, choose one of the three alternatives for the type of keyed match, and move the key variable(s) from the list of Excluded Variables into the bottom list of Key Variables.

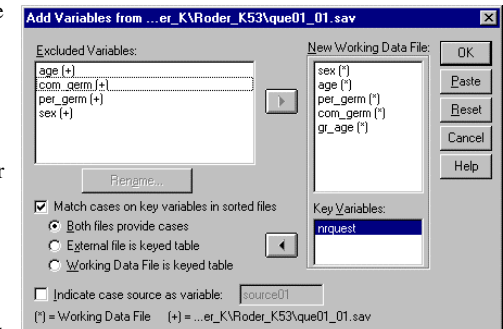


Figure 26: Dialog box for MATCH FILES according to keyed table

Specify the variables for the new working data file:

The list box at the right, *Variables in New Working Data File*, initially contains all the variables in the working data file, marked (*), followed by all differently-named variables in the external file, marked (+). You can add variables to this list, or remove them from it.

The list box at the left, *Excluded Variables*, contains variables that will not normally be included in the new working data file, because they have the same name as variables in the working data file. Initially all of them are from the external file, and are marked (+). Variables marked with (*) are from the (old) working data file. Select one of the following for instructions:

Replacing the values of an existing variable.

Adding a variable with a conflicting name

Excluding a variable.

Using different-named key variables.

Replacing working-file data from external file.

If you select *Indicate case source as variable* a variable is added to the new working data file which has the value 0 for cases that were in the original working data file, and the value 1 for cases that were added from the external file, and may therefore be questionable. (This option is disabled when you select one of the keyed table alternatives, since cases are never added from a keyed table.) To change the name of a variable in the *Excluded Variables* list, select it and press **[Rename]**.

✓ Try yourself to merge SPSS data files. First add cases of two files, then merge files with a different set of variables. Follow the instructions. The resulting syntax file should look similar to the one below.

Example: The MC sample survey

We will apply some more transformation commands on some modified data sets of the MC sample survey. The commands are given in a syntax file, for the beginning the menu options should be used, but try the [Paste] button to see the output in a SPSS syntax file.

```
* Syntax file      :
ex04_1.sps
* Data File charged :
que04_1.sav, que04_2.sav
* Data File saved  :
que04_3.sav
* Data File saved  :
que04_4.sav
* Data File saved  :
que04_5.sav
* Output File      :
```

	nrquest	sex	age	per_germ	com_germ	g
1	4.00	.00	40.00	5.00	1.00	
2	14.00	.00	66.00	2.00	5.00	
3	23.00	.00	27.00	4.00	5.00	

Figure 27: Result of MATCH FILES

```
* Graphics File   :
-----
* Comments        : according to keyed table; File QUE04_5.SAV
* Comments        :
* Author          : CDG-FZ München/Klaus Röder
*****
```

```
GET FILE='C:\Roder_K\Roder_K53\que04_01.sav'.
ADD FILES /FILE=*
/FILE='C:\Roder_K\Roder_K53\que04_02.sav'
/RENAME (per_germ = d0)
/IN=source01
/DROP= d0.
VARIABLE LABELS source01
'Source of observations: C:\Roder_K\Roder_K53\que04_01.sav'.
EXECUTE.
SAVE OUTFILE 'C:\Roder_K\Roder_K53\que04_3.sav'.
```

```
GET FILE='C:\Roder_K\Roder_K53\que04_02.sav'.
EXECUTE.
```

```
MATCH FILES /FILE=*
/FILE='C:\Roder_K\Roder_K53\que04_01.sav'.
/RENAME (age per_germ sex = d0 d1 d2)
/BY nrquest
/DROP= d0 d1 d2.
SAVE OUTFILE 'C:\Roder_K\Roder_K53\que04_4.sav'.
EXECUTE.
```

```
GET
FILE='C:\Roder_K\Roder_K53\que04_02.sav'.
EXECUTE .
MATCH FILES /TABLE=*
/FILE='C:\Roder_K\Roder_K53\que04_01.sav'.
/RENAME (age per_germ sex = d0 d1 d2)
/BY nrquest
/DROP= d0 d1 d2.
SAVE OUTFILE 'C:\Roder_K\Roder_K53\que04_5.sav'.
EXECUTE.
```

Exercise: The Cameroon Household Survey:

In the annex you will find a description of the Cameroon Household Survey. Use the data files CHS96_02.SAV, CHS96_2a.SAV and CHS96_03.SAV, the file containing information about household expenditures.


Use SPSS to merge the data files:

1. add cases of CHS96_02.SAV and CHS96_2a.SAV according to the household identifier BRID and save as CHS96_2b.SAV
2. add variables according to file CHS96_02.SAV and CHS96_03.SAV. All files must be sorted on the household identifier BRID. The new file should contain information about all household members and the expenditure of the household attributed to each of the household members.

Additional exercise:

3. Add new variables according to expenditure per capita and group the household in three income groups (rich, medium ,poor) The new file is called CHS96_04.SAV

5.21. Functions for the Initial Analysis of Data

 In this chapter you will learn to use SPSS functions. There will be very little explanation about the statistical theory behind the functions, existing knowledge is assumed. If you are not sure, what the method does, please consult a statistical text book. The main objective of this chapter is to explain the handling of the functions and the interpretation of the results. You should learn both this in this chapter.

5.22. Frequencies

Menu: [Statistics/Summarize/Frequencies...]

Command Language: *FREQUENCIES*

Move one or more variables into the *Variable(s) list*. A frequency table will be produced for each variable in this list, showing the number of cases for each value.

Select [**Statistics**] to request univariate statistics.

Select [**Charts**] to request bar charts or histograms.

Select [**Format**] to control the order of values in the table, its arrangement on the page, the display of an index, and the display of value labels.

If you deselect the [*Display frequency tables*] box, no frequency table is displayed, but any statistics or charts that you request will appear.

[Statistics]

Quartiles Values that divide the cases into four equal-sized groups.

Mean The arithmetic average; the sum divided by the number of cases.

Cut points for equal groups Values that divide the cases into some number of equal-sized groups

Median A value that is greater than half of the actual data and less than the other half. The 50th percentile.

Percentiles Values above and below which certain percentages of cases fall.

Mode The most commonly occurring value. If several values share the greatest frequency of occurrence, each of them is a mode. The Frequencies procedure reports the smallest of such multiple modes.

Sum The sum or total of the values, across all cases with nonmissing values.

Std. deviation A measure of dispersion around the mean, equal to the square root of the variance. The standard deviation is measured in the same units as the variable itself.

Minimum The smallest value of a numeric variable.

Skewness	A measure of the asymmetry of a distribution. Positive skewness indicates that the more extreme values are greater than the mean; negative skewness indicates that the more extreme values are less than the mean.
Variance	A measure of dispersion around the mean, equal to the sum of squared deviations from the mean divided by one less than the number of cases. The variance is measured in units that are the square of those of the variable itself.
Maximum	The largest value of a numeric variable.
Kurtosis	A measure of the extent to which a distribution is "tail-heavy," compared to a normal distribution. Positive kurtosis indicates more cases in the extreme tails than in a normal distribution with the same variance.
Range	The difference between the largest and smallest values of a numeric variable; the maximum minus the minimum.
S.E. mean	The standard error of the mean of the summarized variable for all the cases in the cell. Available for summarized variables and their totals.

If you select Values are group midpoints, the statistics in Percentile Values and the median statistic are calculated under the assumption that your data have been grouped, and that the actual values in the data are the midpoints of the original groups.

[Charts..]

A **bar chart** displays the frequency count for each value as a separate bar. Bar charts are appropriate for variables measured at the nominal level, or for variables with few distinct values. Values for which the count is zero do not appear. On bar charts, the scale axis can be labeled by *actual frequency counts* or by *percentages*.

A **histogram** displays the frequency count for ranges of values as separate bars. Histograms are appropriate for variables for which it is meaningful to group adjacent values. Space is left for ranges into which no cases fall. A normal curve can be superimposed on a histogram as an aid in judging whether the variable is normally distributed.

(You can obtain a group of histograms that all use the same scale by selecting from the [Statistics/Summarize/Explore..] submenu, or by using the SPSS command language.)

[Format..]

The frequency table can be arranged according to the actual values in the data, or according to the count (frequency of occurrence) of those values, and in either ascending or descending order. However, if you request a histogram (in [Charts]) or you request percentiles or equal cut points (in [Statistics]), the table is always arranged in ascending order of values.

[Suppress tables with more than] a specified number of values displays the remaining tables in standard format.

We want to know about the distribution of age and age groups, as well as about the distribution to the answers concerning the most popular German personality and best known German company.



Example: The MC sample survey

We will apply the frequencies function to the data of the MC sample survey. The commands are given in a syntax file, the results are displayed in an output file and in a graphics file.

```
*****
* Syntax file           : ex05_1.sps
* Data File charged    : que_01_3.sav
* Output File          : res05_1.lst
* Graphics File        : gra05_1.jpg, gra05_2.jpg, gra05_3.jpg
* Comments             : Syntax and Output Files are displayed
* Comments             : in one file
* Comments             : The different section are marked by
* Comments             : the respective headers
* Author               : CDG-FZ München/klaus röder
*****
* Syntax File
GET
FILE='C:\Roder_K\Roder_K53\Quest03.sav'.
EXECUTE .
FREQUENCIES
  VARIABLES=age per_germ com_germ
  /STATISTICS=STDDEV VARIANCE RANGE MINIMUM MAXIMUM SEMEAN MEAN
SKEWNESS
  SESKW KURTOSIS SEKURT .
FREQUENCIES
  VARIABLES=age
  /HISTOGRAM NORMAL.
FREQUENCIES
  VARIABLES=per_germ com_germ
  /BARCHART FREQ
  /ORDER ANALYSIS .
```

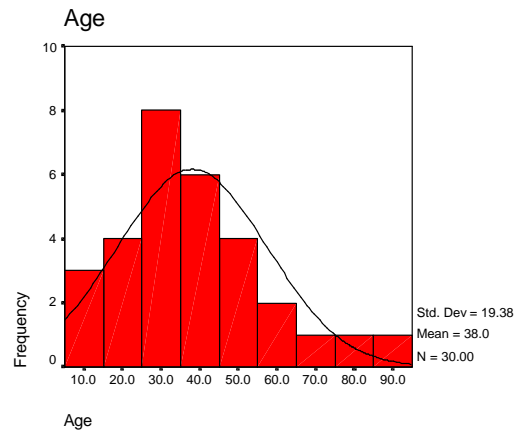

Selected output:

Frequencies

Statistics

	Age	Best known German Personality	Best known German Company
N	Valid 30	30	30
	Missing 0	0	0
Mean	38.0333	3.5667	3.4667
Std. Error of Mean	3.5375	.3276	.3706
Std. Deviation	19.3756	1.7943	2.0297
Variance	375.4126	3.2195	4.1195
Skewness	.811	-.174	.158
Std. Error of Skewness	.427	.427	.427
Kurtosis	.374	-1.350	-1.723
Std. Error of Kurtosis	.833	.833	.833

Histogram



Age	Age	Best known German Personality	Best known German Company
Range	79.00	5.00	5.00
Minimum	9.00	1.00	1.00
Maximum	88.00	6.00	6.00

Frequency Table

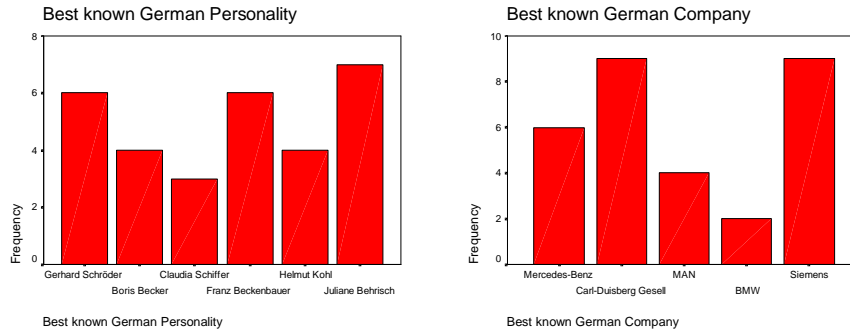
Best known German Personality

Valid	Frequency	Percent	Valid Percent	Cumulative Percent
Gerhard Schröder	6	20.0	20.0	20.0
Boris Becker	4	13.3	13.3	33.3
Claudia Schiffer	3	10.0	10.0	43.3
Franz Beckenbauer	6	20.0	20.0	63.3
Helmut Kohl	4	13.3	13.3	76.7
Steffi Graf	7	23.3	23.3	100.0
Total	30	100.0	100.0	

Best known German Company

Valid	Frequency	Percent	Valid Percent	Cumulative Percent
Mercedes-Benz	6	20.0	20.0	20.0
Carl-Duisberg Gesellschaft	9	30.0	30.0	50.0
MAN	4	13.3	13.3	63.3
BMW	2	6.7	6.7	70.0
Siemens	9	30.0	30.0	100.0
Total	30	100.0	100.0	

Bar Chart



As a result you observe the following:

The distribution of age differs from the normal distribution as far as the respondents seem to be somewhat younger than a normal distribution would let you expect (what about the Skewness and Kurtosis?, they are both slightly positive). From the frequency tables you will find out, that Julia Behrisch rivals with some politicians for being the most popular personality and the companies Siemens and Carl-Duisberg Gesellschaft are the best known ones.

5.23. Explorative data analysis

Like the *Frequencies* function delivers some general statistical information about the data set, the function *Explore* allows to observe statistics and listings about all observations as well as groups of observations. It incorporates some graphical techniques which are especially useful for a first analysis of the data set, comparison among groups of responses and test of normality as well as explicative graphics (like boxplots and stem and leaf graphics) or a comparison of variances. Like the other functions in this chapter, this is a first approach to examine or explore the data, especially suited for a first glance at a bulky data set.

Menu: [Statistics, Summarize, Explore]
Command language: *EXAMINE*

Move at least one variable into the Dependent List. Statistics and plots from this command will describe the distribution(s) of the variables in this list. You may move one or more grouping variables into the Factor List. The values of factor variables define groups of cases, which will be described separately.

You may move a variable into the Label Cases by box. This variable is used to identify unusual cases in some of the output. Select one of the alternatives in the Display group to choose plots, statistics or both for the output. These alternatives activate the **[Plots]** and **[Statistics]** pushbuttons:

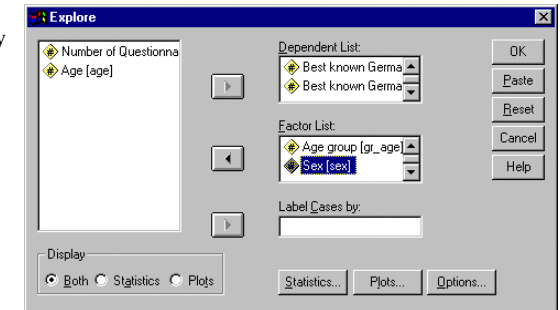


Figure 28: 'Explore' Dialog Box

[Plots] allows you to control the output of plots.

[Statistics] allows you to request descriptive statistics and a listing of cases with the 5 smallest and the 5 largest values.

[Options] allows you to determine the treatment of missing values.

Example: The MC sample survey

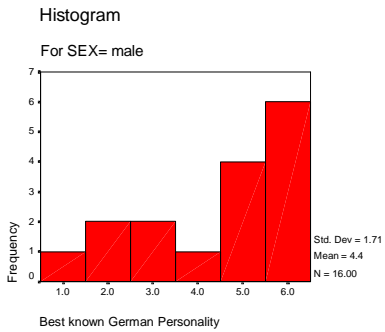
We will try to get more information about the distribution of sex and age groups from our sample data.

We will apply the explore function to the data of the MC sample survey. The commands are given in a syntax file, the results are displayed in an output file and in a graphics file.

```
*****
* Syntax file       : ex05_2.sps
* Data File charged : quest3.sav
```

```
* Output File      : res_05_2.lst
* Graphics File    : gra_05_3.jpg,gra_05_4.jpg
* Comments         : Syntax and Output Files are displayed
* Comments         : in one file
* Comments         : The different section are marked by
* Comments         : the respective headers
* Author           : CDG-FZ München/klaus röder
*****
* Syntax File
GET
FILE='C:\Roder_K\Roder_K53\Quest03.sav'.
EXECUTE .
EXAMINE
VARIABLES=per_germ com_germ BY gr_age sex
/PLOT BOXPLOT STEMLEAF HISTOGRAM NPLOT SPREADLEVEL
/COMPARE GROUP
/STATISTICS NONE
/CINTERVAL 95
/MISSING LISTWISE
/NOTOTAL.
EXAMINE
VARIABLES=per_germ com_germ age
/PLOT NONE
/STATISTICS DESCRIPTIVES
/CINTERVAL 95
/MISSING LISTWISE
/NOTOTAL.
```

*** Results (partial):. Histograms**



Stem-and-Leaf Plots

Best known German Personality
Stem-and-Leaf Plot for
SEX= male

Frequency	Stem & Leaf
1.00	1 . 0
2.00	2 . 00
2.00	3 . 00
1.00	4 . 0
4.00	5 . 0000
6.00	6 . 000000

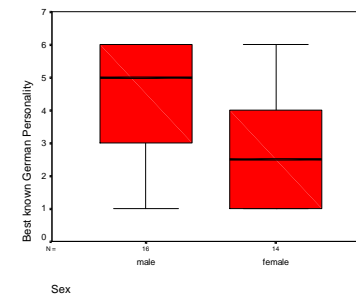
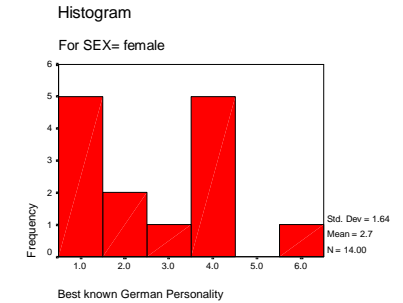
Stem width: 1.00
Each leaf: 1 case(s)

Best known German Personality Stem-and-Leaf Plot for
SEX= female

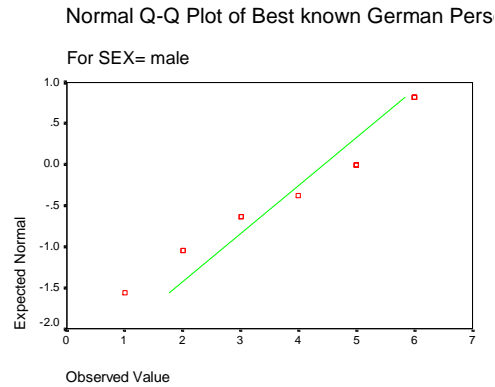
Frequency	Stem & Leaf
5.00	1 . 00000
2.00	2 . 00
1.00	3 . 0
5.00	4 . 00000
.00	5 .
1.00	6 . 0

Stem width: 1.00
Each leaf: 1 case(s)

Boxplots



Normal Q-Q Plots



Before starting the analysis, let us explain briefly the different plot and display techniques used:

Boxplots

Summary plots based on the median, quartiles, and extreme values. Boxplots are formed from "boxes", which contain the 50% of values falling between the 25th and 75th percentiles, and the "whiskers", lines that extend from the box to the highest and lowest values, excluding outliers. An asterisk indicates the median. Boxplots also convey information about spread and skewness. SPSS boxplot output also identifies outliers and extreme values.

Stem-and-leaf plots

Summary plots that are similar to histograms, but preserve more information. Rather than using the same symbol to represent all cases, stem-and-leaf plots use actual data values. Each observed value is divided into two components: a stem (leading digits) and a leaf (trailing digit). The leaves determine the plot symbols.

Normal Q-Q plots

plots the quantiles of a variable's distribution against the quantiles of the normal distribution

Analysis shows, that age groups differ in responses about the best known personality as do different sexes. For the last analysis, we can reject the hypothesis, that data are normal distributed because the level of significance of the K-S Test is .0965. The Normal Q-Q Plot shows the distribution (actual and normal) in a graph. This functions allows for the first time **bivariate** statistics: statistical relation between two variables as opposed to **univariate** statistics: statistics about one variable only.

5.24. Crosstabs

Crosstabs is used to count the number of cases that have different combinations of values of two or more variables, and to calculate summary statistics and tests. For example, if you want to know if there is a relationship between the region of the city in which a person lives and the person's willingness to buy your product, you can make a table that shows how many people say they would buy your product, and how many say they would not, for each of the regions. Or you can look at the relationship between job satisfaction and job performance ratings separately for males and females. The variables you use to form the categories within which the counts are obtained should have a limited number of distinct values. For example, if you want to look at the relationship between salary and job satisfaction, and you have recorded salary to the nearest dollar, you should use the Recode facility to group salaries into ranges, such as less than \$20,000, \$20,000 to \$30,000, and so on.

Menu: Statistics/Summarize/Crosstabs...
Command language: *CROSSTABS*

Move one or more variables into the Row(s) box, and one or more variables into the Column(s) box. Categories of a row variable form the rows of the table, and categories of a column variable form the columns. A table is produced for each row variable in combination with each column variable. Optional **Layers**, discussed below, are used to specify control variables.

[Statistics] allows you to request bivariate statistics relating the row and column variables.

[Cells] allows you to control the contents of the cells making up the table.

[Format] allows you to control the labeling and appearance of the tables.

Select *Suppress Tables* to see statistics only, without the actual tables.

Layers:

Table(s) defined by the row and column variables are produced within each category of the variable(s) in the first layer. If there is a second layer, tables are produced within each combination of categories for each 1st-layer variable with each 2nd-layer variable; and so on.

Example: The MC sample survey

We want to know more about the relation between age groups and replies to the questions about the most popular German personality and best known company.



We will apply the crosstabs function to the data of the MC sample survey. The commands are given in a syntax file, the results are partially displayed in an output file.

```
*****
* Syntax file           : ex05_3.sps
* Data File charged    : quest03.sav
* Output File          : res_05_3.lst
```

```
* Graphics File      :
* Comments          : Syntax and Output Files are displayed
* Comments          : in one file
* Comments          : The different section are marked by
* Comments          : the respective headers
* Author            : CDG-FZ München/klaus röder
*****
* Syntax File
GET
FILE='C:\Roder_K\Roder_K53\Quest03.sav'.
EXECUTE .
COMPUTE gr_age2=0.
IF (gr_age < 3) gr_age2 = 1 .
IF (gr_age >=3 & gr_age <5 )gr_age2 = 2.
IF (gr_age >=5 & gr_age <6 )gr_age2 = 3 .
IF (gr_age >=6) gr_age2 = 4 .
EXECUTE .
CROSSTABS
/TABLES=per_germ BY gr_age2
/FORMAT= AVALUE NOINDEX BOX LABELS TABLES
/STATISTIC=CHISQ CORR
/CELLS= COUNT .
EXECUTE .
CROSSTABS
/TABLES=per_germ BY gr_age2 BY sex
/FORMAT= AVALUE NOINDEX BOX LABELS TABLES
/CELLS= COUNT .
SAVE OUTFILE='C:\SPSSWINE\QUE_05_1.SAV'
/COMPRESSED.
```

* Results (partial)

Crosstabs

Best known German Personality * GR_AGE2 Crosstabulation
Count

	GR_AGE2					Total
		1.00	2.00	3.00	4.00	
Best known German Personality	Gerhard Schröder	1	3		2	6
	Boris Becker	1		3		4
	Claudia Schiffer	1			2	3
	Franz Beckenbauer	1	4	1		6
	Helmut Kohl		2		2	4
	Steffi Graf	1	5		1	7
Total		5	14	4	7	30

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	26.676	15	.031
Likelihood Ratio	28.192	15	.020
Linear-by-Linear Association	.251	1	.617
N of Valid Cases	30		

a 24 cells (100.0%) have expected count less than 5. The minimum expected count is .40.

Symmetric Measures

	Value	Asymp. Std. Error	Approx. T	Approx. Sig.
Interval by Interval	Pearson's R .093	.181	-.494	.625
Ordinal by Ordinal	Spearman Correlation .093	.183	-.497	.623
N of Valid Cases	30			

a Not assuming the null hypothesis.

b Using the asymptotic standard error assuming the null hypothesis.

c Based on normal approximation.

Best known German Personality * GR_AGE2 * Sex Crosstabulation
Count

Sex	GR_AGE2	Total				
		1.00	2.00	3.00	4.00	
male	Best known German Personality	1				1
	Gerhard Schröder			2		2
	Boris Becker					2
	Claudia Schiffer					1
	Franz Beckenbauer		2			4
	Helmut Kohl	1	4			6
	Steffi Graf	1	7	3	5	16
female	Total	1	7	3	5	16
	Best known German Personality	1	2			5
	Gerhard Schröder			1		2
	Boris Becker	1				2
	Claudia Schiffer	1				1

	Franz	1	4		5
	Beckenbauer				
	Steffi Graf		1		1
Total		4	7	1	2
					14

Before starting the analysis, let us explain briefly some indicators and coefficients:

Chi-Square (Crosstabs)

Statistic used to test the hypothesis that the row and column variables are independent. Should not be used if any cell has an expected value less than 1, or if more than 20% of the cells have expected values less than 5 (which is the case in our small example). The Pearson chi-square is the most widely used form, while the likelihood-ratio chi-square is based on maximum-likelihood theory.

The Mantel-Haenszel chi-square, a measure of association for ordinal variables, is also displayed by this option. This coefficient would be more appropriate because *per_germ* is a ordinal variable.

Although the hypothesis that the variables are independent can be rejected, because both coefficients are far too low (Pearson = .03 and Mantel-Haenszel = .52637). In other words, the probability of independence according to Pearson is 3%.

The second crosstabs command without statistics is calculated with a layer variable (*sex*), the results for female respondents are displayed.


Exercise: The Cameroon Household Survey:

In the annex you will find a description of the Cameroon Household Survey. Use the data file CHS96_04.SAV

Use the SPSS functions to analyze the data:

1. Count the frequency of the *Highest School Degree obtained*, exclude minors (< 18 years)
2. Count the frequency of public sector employees and private sector employees.
3. Create a new age group variable (4 values only). Analyze by explorative data analysis the relations between age groups, sex and the variables: *Ever Attended School* and *Highest School Degree obtained*
4. Analyze the independence of the above mentioned variables and sex.
5. Analyze the independence of *medical expenses* and the above mentioned variables.
6. Display these variables of exercise 4. in crosstabs. Display the results for age groups and *medical expenses* and age groups and *educational expenses*. Do a second crosstabs commands with layers for the sex of the head of household.
Take the primary occupation into account only!

5.25. Syntax Files

 You should learn the essentials about syntax files and be familiar to use the command interface of SPSS as well as the menu driven interface.

5.26. Syntax Files: General considerations

Most commands can be executed in SPSS in two ways. The simple way is to call a menu and the subsequent dialogue boxes to perform the command. The other is to write a command in a syntax file and execute it. This may seem to be more complicated because you have to learn the command syntax but there are multiple advantages:

- You can store and repeat the set of commands
- Repeat the commands, paste and cut part of the command sequence
- Use the files for later documentation
- Some commands or command options can only be executed by commands in a syntax file
- You can use command sequences of earlier SPSS versions and use them again with later versions.

Create, Open, Save and Close Syntax Files

Menu: [File/New/Syntax]

this opens a text window with controls designed to help you create and submit SPSS command syntax. The Paste button in dialogue boxes pastes command syntax into the designated syntax window.

Menu: [File/Open/Syntax]

Select or enter the name of the text file to open in a Syntax or Output Window. The file list lets you choose the file to open. The Directories list allows you to choose a directory. You can select the type of file that you want to open from the File Type list. You can select a drive by choosing one of the options from the Drives list. You can also type the fully-qualified filename into the File Name box.

Menu: [File/Save]

Saves the file in the active window. If the active window contains a data, chart, or text file (like a syntax file), it is saved using its existing name. If the file has not been named this is the same as Save As. If the same filename is in two windows, you should use Save As or Close instead of Save, to avoid the possibility of overwriting one file with another.

If you save a text file (from an output or syntax window) that already exists on the disk, the older version is renamed to have the extension bak.

Menu: [File/Close]

Closes the file in the active window. Alternatively use [Alt] + [F4] or the WINDOWS Icon.

5.27. Editing Text in Output and Syntax Windows

Edit the text in an output window or a syntax window using either a mouse or the keyboard: Characters that you type appear at the insertion point, which you can move with the mouse or with the movement keys on the keyboard.

Select text with the left mouse button (for entire lines) or the right mouse button (for portions of a line, or for rectangles). The left button is preferable when copying text to other applications. The F8 keyboard accelerator (Extend Selection) lets you sweep out a selection with cursor movement keys. Other accelerators are available for selecting text.

Selecting Text from the Keyboard


Function	Key
Select text to beginning of document	Shift-Ctrl-Home
Select text to end of document	Shift-Ctrl-End
Select text left	Shift-Left Arrow
Select word left	Shift-Ctrl-Left Arrow
Select text right	Shift-Right Arrow
Select word right	Shift-Ctrl-Right Arrow
Select text to beginning of line	Shift-Home
Select text to end of line	Shift-End


The Edit menu provides Cut, Copy, Paste, Search, and Replace commands. These commands also have accelerators.

Using Editing Commands from the Keyboard

Function	Key
Search	F5
Replace	Shift-F5
Edit Copy Selection	Ctrl-Insert or Ctrl-C
Edit Cut Selection	Shift-Delete or Ctrl-X
Edit Paste Selection	Shift-Insert or Ctrl-V
Delete selection	Delete
Define a term (in output window)	Ctrl-G
Round numbers within selected area	Ctrl-0 (Ctrl-zero)

5.28. Execute commands in Syntax Windows

The **[Run]** or  button on a syntax window executes the command(s) containing the selection. To run a group of commands, select them with the mouse or keyboard before selecting **[Run]**.

Where  significant functionality is available only through the command language, a Help topic for the dialog box will list these features. Syntax charts are also available on-line. For complete information about the SPSS command language, refer to the SPSS Reference Guide.

5.29. The command syntax

Each SPSS command is described in a Syntax Diagram. Call the help feature to show a list of all Diagrams.

Interpreting Syntax Diagrams

Note: Lines in a syntax diagram do not "wrap" on your screen. If the diagram is too wide for the current Help window, a scroll bar appears at the bottom of the window. You may wish to make the Help window short and wide, so that syntax diagrams are visible without scrolling.

- Each command terminates with a period. The command terminator (a period) is not shown in the syntax diagram.
- Elements shown in capital letters are keywords. You must enter them as shown, although you can abbreviate them as long as there is no ambiguity. Elements in lower case describe specifications you supply.
- Elements shown in [square brackets] are optional.
- Parentheses, apostrophes, and quotation marks are required where shown, unless they are in square brackets. A pair of apostrophes is generally equivalent to a pair of quotation marks.
- Braces {} indicate a choice between elements.
- The word varlist stands for a list of variable names.

Example: The MC sample survey



We will apply one command to be written and edited in a command window.

This command is **DESCRIPTIVES**

This is the command syntax from the Syntax Diagram

```
DESCRIPTIVES [VARIABLES=] varname[(zname)] [varname...]
[/MISSING={VARIABLE**} [INCLUDE]]
           {LISTWISE }
[/SAVE]
[/STATISTICS={DEFAULT**}[MEAN**][MIN**][SKEWNESS]]
           [STDDEV**][SEMEAN][MAX**][KURTOSIS]
           [VARIANCE][SUM][RANGE][ALL]
[/SORT={ {MEAN } } [ { (A) } ] ]
           {SMEAN } { (D) }
           {STDDEV }
           {VARIANCE}
           {KURTOSIS}
           {SKEWNESS}
           {RANGE }
           {MIN }
           {MAX }
           {SUM }
           {NAME }
```

**Default if the subcommand is omitted.

This is the syntax file to be edited to call the command

```
*****
* Syntax file           : ex06_1.sps
* Data File charged    : quest04.sav
* Output File          : -----
* Graphics File        : -----
* Comments              :
* Comments              :
* Author                : CDG-FZ München/klaus röder
*****
* Syntax File
GET
  FILE='C:\Roder_K\Roder_K53\Quest04.sav'.
EXECUTE .
DESCRIPTIVES
  VARIABLES=gr_age per_germ com_germ
  /FORMAT=LABELS INDEX
  /STATISTICS=MEAN STDDEV VARIANCE RANGE MIN MAX
  /SORT=MEAN (A) .
```

Exercise: The Cameroon Household Survey:

In the annex you will find a description of the Cameroon Household Survey. Use the data file CHS96_04.SAV


Use SPSS a Syntax Diagram (you might use the **[Paste]** button to help out) to show the Variance, Kurtosis and Skewness of the following Variables, use *DESCRIPTIVES* for

1. *Age*
2. *Highest School Degree obtained*

Use SPSS a Syntax Diagram (you might use the **[Paste]** button to help out) to show

3. crosstabs between *Employment Status* and *Expenditure on Health*

5.30. The Presentation of Data

 The goal of statistical analysis is to produce results. These results have to be communicated to the users of statistics, sometime these are statisticians, sometimes they are not. So the presentation has to be comprehensive, informative and attractive. This means that the statistician has to produce results and present them according to his or her audience. The explanations necessary to understand results will always be better. Graphics and charts should always be preferred as means of presentation even as a supplement to more traditional presentation. This chapter deals with the possibilities SPSS offers to publish and display statistical data, listing, tabulation and graphical representation. This is in brief the techniques you will use for most statistical publication. You should learn these methods and functions, you should know when and how to use them.

5.31. Listings and Reports of Data

5.31.1. List Cases

displays the values of variables for cases in the data file.

Menu: *no Menu available*

Command Language: *List*

Command Syntax:

```
LIST [[VARIABLES=]{ALL** }]  
      {varlist}  
[/FORMAT=[{WRAP**}] [{UNNUMBERED**}]]  
      {SINGLE} {NUMBERED }  
[/CASES=[FROM {1**}][TO {eof**}][BY {1**}]]  
      {n } {n } {n }
```

Move one or more variables into the Variable(s) list. The values of the selected variables will be displayed in the listing file. Select one of the alternatives in the *CASES TO* group to choose the cases to be listed. To list a subset of cases beginning with the first case, select *1 TO* and then enter the number of the last case to be listed. Interval allows you to determine the increment used for choosing cases.

Select *FORMAT NUMBERED* if you want to include the sequence number of each case in the listing.

5.31.2. Report Summaries in Rows

produces reports in which different summary statistics are laid out in rows. Case listings are also available from this command, with or without summary statistics.

Menu: [Statistics/Summarize/Report Summaries in Rows]
Command Language: *Report*

Minimum specification: You must select either the *Display cases* option, or at least one summary statistic (from either of the Summary subdialogs).

The data columns in a report contain statistics summarizing your variables. Break columns (at the left) display individual values of grouping variables, to identify the group statistics in the data columns. You may wish to look at an example of a report with summaries in rows, to compare it with a listing report, with summaries in rows, which also includes rows containing values of individual cases; and also with a report with summaries in columns. (The latter is a separate command on the menu.)

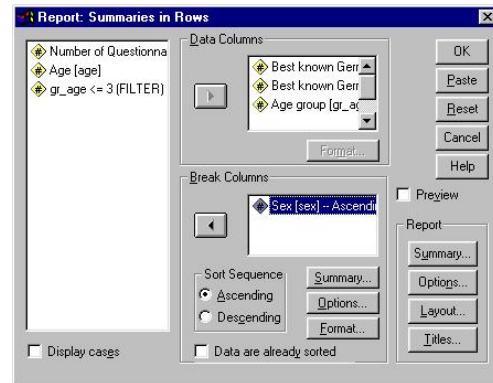


Figure 29: Report Summaries in Rows

Move at least one variable into the Data Columns list. One column in the report is generated for each variable in the Data Columns box. It contains any summary statistics you request for the variable, and (if requested) the actual value or value label for each case. To specify an output format for a variable, select that variable in the Data Columns box and then select **[Format]**.

Break Columns define subgroups of cases for summary statistics in the report display. Individual values of each break variable appear, sorted, in a separate column to the left of all the data columns. Select one or more categorical variables as break variables, and move them into the Break Columns box, in their desired order.

You can sort the categories of break variables either in ascending or descending order. Select the variable in the Break Columns box and then select either *Ascending* or *Descending* from the Sort Sequence group. The sort order for each variable appears next to the variable name in the Break Columns box as (A) or (D). String variables are sorted in forward or reverse alphabetical order.

If the data are already sorted, select *Data are already sorted* so that SPSS will not sort them again. This is often useful as you refine a report format. The first report generated (with a

given set of break variables) sorts the file, so later reports can be produced faster if you select this option.

Pushbuttons in the Break Columns group determine what is displayed at the end of the subgroup of cases defined by each category of a break variable. Select a break variable, and then select:

[Summary] to request univariate statistics on data-column variables for each subgroup of cases.

[Options] to control page breaks and blank lines before the subgroup statistics.

[Format] to control display of the values or value labels in the break column.

Select *Display cases* to display the actual values (or value labels) of the data-column variables for every case. This produces a listing report, which can be much longer than a summary report.

Select *Preview* to see a sample page of the report according to the current specifications in this dialog box. The complete report is not generated if this option is selected.

The Report group allows you to request overall statistics and to control page breaks, page layout, and titles:

Select **[Summary]** to request univariate statistics on data-column variables for all cases.

Select **[Options]** to control treatment of missing values, data sorting, and page numbers.

Select **[Layout]** to customize the page layout of the report.

Select **[Titles]** to specify header and footer text for the top and bottom of each page of the report.

5.31.3. Report Summaries in Columns

produces reports in which different summary statistics are laid out in separate columns.

Menu:[Statistics/Summarize/Summaries in Columns]
Command Language: *Report*

The data columns in a report contain statistics summarizing your variables. Break columns (at the left) display individual values of grouping variables, to identify the group statistics in the data columns. You may wish to look at an example of a report with summaries in columns, to compare with one displaying summaries in rows.

Move at least one variable into the Data Columns list. It will initially appear with the Sum statistic. (To obtain a different statistic, select the variable:statistic pair in the Data Columns box and click [Summary].) One column in the report is generated for each variable:statistic pair in the Data Columns box, containing the indicated summary statistics for the variable, calculated on the cases in each break group.

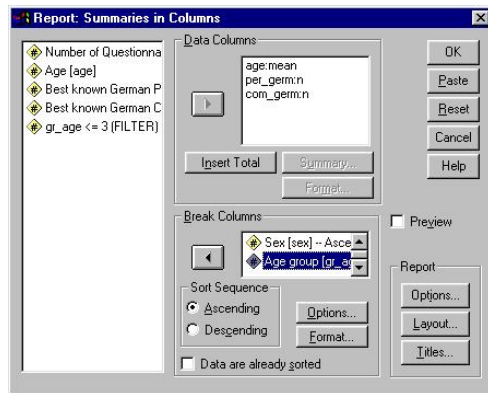


Figure 30: Report Summaries in Columns

The Data Columns group also contains a set of buttons that allow you to control the format and content of the data columns:

To display a column containing the sum, mean, ratio, or other function of existing columns, select [Insert Total]. This places a "variable" called Total into the Data Columns list, below the variable currently selected in the Data Columns list. With Total selected in the Data Columns box, select [Summary] to indicate how the total column should be computed.

To change the statistic in a variable:statistic pair in the Data Columns list, select the pair and then select [Summary].

To specify an output format for a variable or total, select that variable or total in the Data Columns list and then select [Format].

Break Columns define subgroups of cases for summary statistics in the report display.

Individual values of each break variable appear, sorted, in a separate column to the left of all the data columns. Select one or more categorical variables as break variables, and move them into the Break Columns box, in their desired order.

You can sort the categories of break variables either in ascending or descending order. Select the variable in the Break Columns box and then select either Ascending or Descending from the Sort Sequence group. The sort order for each variable appears next to the variable name in the Break Columns box as (A) or (D). String variables are sorted in forward or reverse alphabetical order.

If the data are already sorted, select Data are already sorted so that SPSS will not sort them again. This is often useful as you refine a report format. The first report generated (with a given set of break variables) sorts the file, so later reports can be produced faster if you select this option.

Pushbuttons in the Break Columns group affect the spacing and appearance of the values displayed in the break column(s). Select a variable in the Break Columns list and then select:

[Options] to control page breaks and blank lines before the subtotals.

[Format] to control display of the values or value labels in the break column.

Select Preview to see a sample page of the report according to the current specifications in this dialog box. The complete report is not generated if this option is selected.

The Report group allows you to request overall statistics and to control page breaks, page layout, and titles:

Select Options to control treatment of missing values, data sorting, and page numbers.

Select Layout to customize the page layout of the report.

Select Titles to specify header and footer text for the top and bottom of each page of the report.

Example: The MC sample survey



We will apply the above mentioned commands to the MC sample survey .

```
*****
* Syntax file           : ex07_1.sps
* Data File charged    : quest04.sav
* Output File          : res_07_1.lst
* Graphics File        :
* Comments              : Syntax and Output Files are displayed
* Comments              : in one file
* Comments              : The different section are marked by
* Comments              : the respective headers
* Author                : CDG-FZ München/klaus röder
*****
```

```
GET
  FILE='C:\Roder_K\Roder_K53\Quest04.sav'.
EXECUTE .
```

```
LIST
  VARIABLES=gr_age age per_germ com_germ
  /CASES= FROM 1 TO 10 BY 1
  /FORMAT= WRAP UNNUMBERED .

SORT CASES BY sex (A) .
Report
  /FORMAT= CHWRAP(ON) PREVIEW(OFF) CHALIGN(BOTTOM)
  UNDERSCORE(ON)
  ONEBREAKCOLUMN(OFF) CHDSPACE(1) SUMSPACE(0) AUTOMATIC LIST
  BRKSPACE(-1)
  PAGE(1) MISSING'.' LENGTH(1, 59) ALIGN(LEFT) TSPACE(1)
  FTSPACE(1)
```

```
MARGINS(1,74)
/TITLE=
  RIGHT 'Page )PAGE'
/VARIABLES
per_germ (VALUES) (RIGHT) (OFFSET(0)) (17)
com_germ (VALUES) (RIGHT) (OFFSET(0)) (26)
gr_age (VALUES) (RIGHT) (OFFSET(0)) (11)
/BREAK sex (LABELS) (LEFT) (OFFSET(0)) (13) .
SORT CASES BY sex (A) gr_age (A) .
Report
  /FORMAT= CHWRAP(ON) BRKSPACE(-1) SUMSPACE(0) AUTOMATIC
  PREVIEW(OFF) CHALIGN(BOTTOM) CHDSPACE(1)
  UNDERScore(ON) ONEBREAKCOL(OFF)
  PAGE(1) MISSING'.' LENGTH(1, 59)ALIGN(LEFT) TSPACE(1)
FTSPACE(1)
MARGINS(1,71)
/TITLE=
  RIGHT 'Page )PAGE'
/VARIABLES
age 'Age' 'Mean' (RIGHT) (OFFSET(0))(13)
per_germ 'Best known German Personality' 'N' (RIGHT)
(OFFSET(0)) (13)
com_germ 'Best known German Company' 'N' (RIGHT)
(OFFSET(0)) (13)
/BREAK sex (LABELS) (LEFT) (OFFSET(0)) (SKIP(1))(13)
/BREAK gr_age (LABELS) (LEFT) (OFFSET(0)) (SKIP(1))(11)
/SUMMARY MEAN(age) SKIP(0) VALIDN(per_germ)
VALIDN(com_germ) .
```

* Results

* List cases

List

GR_AGE	AGE	PER_GERM	COM_GERM
1.00	9.00	6.00	2.00
2.00	13.00	2.00	2.00
2.00	14.00	3.00	2.00
2.00	15.00	4.00	2.00
2.00	18.00	1.00	4.00
3.00	22.00	5.00	6.00
3.00	23.00	1.00	1.00
3.00	26.00	4.00	2.00
3.00	26.00	4.00	6.00
3.00	27.00	5.00	6.00

Number of cases read: 10 Number of cases listed: 10

* Report Summaries in Rows

Sex	Best known German Personality	Best known German Company	Age group
male	6.00	2.00	1.00
	5.00	6.00	3.00
	1.00	1.00	3.00
	5.00	6.00	3.00
	5.00	6.00	3.00
female	2.00	2.00	2.00
	3.00	2.00	2.00
	4.00	2.00	2.00
	1.00	4.00	2.00
	4.00	2.00	3.00
	4.00	6.00	3.00
	6.00	5.00	3.00

* Report Summaries in Columns

Sex	Age group	Age Mean	Best known German Personality N	Best known German Company N
male	<= 10	9.00	1	1
	20< <=30	24.75	4	4
female	10< <=20	15.00	4	4
	20< <=30	27.00	3	3

5.32. The Tables Module

The supplementary module Tables is a combination of list functions, report functions, crosstabs and descriptive statistics. All these SPSS functions merge into a set of commands which allow

to produce tables almost ready for publication. Without any problems these tables produced can be incorporated in any report using the ordinary data exchange opportunities WINDOWS offers.

5.32.1. Basic Tables

produces publication-quality tables displaying crosstabulations and subgroup statistics. It allows nested layouts. The same statistics are reported for all variables summarized in the table.

Menu : [Statistics/Custom
Tables/Basic Tables]
Command Language: *TABLES*

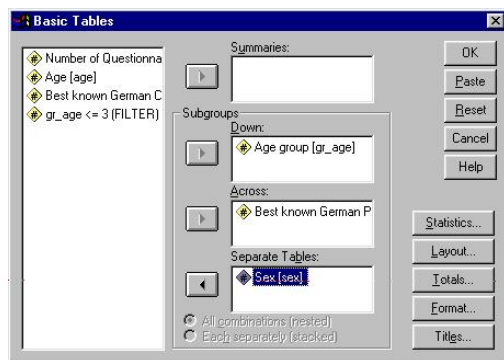


Figure 31: Basic Tables

Move one or more variables into the Summaries list, or into any of the three Subgroups lists, or both. At minimum, you need only one variable in any one of these boxes. If you specify a variable for Summaries, you should select **[Statistics]** to indicate which statistic(s) you want to see.


To obtain simple frequency tables or crosstabulations containing counts or percentages, leave the Summaries list empty and move grouping variable(s) into the Subgroups lists.

To obtain subgroup statistics, move the variable(s) for which you want statistics into the Summaries list and the grouping variable(s) into the Subgroups lists. Select **[Statistics]** to specify the desired statistics. Be sure to select appropriate statistics for the summarized variables. To obtain overall statistics, move the variable(s) for which you want statistics into the Summaries list and select **[Statistics]** to specify the desired statistics. Be sure to select statistics that are appropriate for the summarized variables.

The three Subgroups lists let you define cells to display subgroups of cases in any combination of:

Down the page (as separate rows).

Across the page (as separate columns). The combination of Down and Across generates a crosstabular display.

(Separate Tables) Spread across Separate Pages. Only one layer of the table is visible at a time. You can view other layers after the table is displayed in the Viewer by double-clicking the table and clicking the arrows on the layer pivot icon. 

When you specify more than one grouping variable in any one dimension, you can choose between the alternatives *All combinations (nested)* and *Each separately (stacked)*. (The General Tables command offers more detailed control over the relation between multiple grouping variables in a dimension.)

Select **[Statistics]** to specify the statistics in the table. For statistics on summary variables, select the variable in the Summaries list first. By default the means of summary variables are displayed if there are any summary variables, or the case counts if there are none.

Select **[Layout]** to specify the summary dimension, the statistics dimension, the relative position of summary variables and subgroups in the summary dimension, and to suppress variable labels for grouping variables.

If you have entered a variable into any of the Subgroups lists, you can select **[Totals]** to request total statistics (after the subgroup statistics) for grouping variable(s).

Select **[Format]** to specify margins, column sizes, borders, the appearance of empty cells and undefined statistics, and the characters at which labels can split between lines.

Select **[Titles]** to specify lines of text for the table header, footer, and top-left corner.

Example: The MC sample survey

We will apply the above mentioned Tables commands to the MC sample survey . 

```
*****
* Syntax file           : ex07_2.sps
* Data File charged    : quest05.sav
* Output File          : res_07_2.lst
* Graphics File        :
* Comments              : Syntax and Output Files are displayed
* Comments              : in one file
* Comments              : The different section are marked by
* Comments              : the respective headers
* Author                : CDG-FZ München/klaus röder
*****
GET
  FILE='C:\Roder_K\Roder_K53\Quest05.sav'.
EXECUTE .
* Basic Tables.
TABLES
  /FORMAT BLANK MISSING('.')
  /TABLES gr_age
  BY per_germ > (STATISTICS)
  BY sex.
```

* Results

Tables

(only results for male respondents are displayed)
Sex male

Age group		Gerhard Schröder	Boris Becker	Claudia Schiffer	Franz Beckenbauer	Helmut Kohl	Steffi Graf
<= 10							1
10< <=20							
20< <=30		1				2	1
30< <=40							3
40< <=50			2		1		
50< <=60						1	1
60< <=70				2		1	
70<							

5.32.2. General Tables

provides the greatest amount of control over the content and format of tables, at the cost of somewhat more complex specifications. Different statistics can be displayed for different variables. The General Tables command allows you to nest some variables and stack others. It also supports tables involving multiple-response and multiple-dichotomy sets.

Menu: [Statistics/Custom Tables/General Tables]

Command Language: *TABLES*

Specify where Statistics Labels Appear to determine where the labels for cell statistics appear. (These alternatives are sometimes unavailable, when only a single location is possible.)

Move one or more variables into any or all of the Row list, the Column list, or the Layer list. You can use either the regular variables in the source list at the upper left, or you can use Mult Response sets. (To add multiple-response or multiple-dichotomy sets to the Mult Response list for use in the table, select [Mult Response Sets].)

In order for the [OK] and [Paste] buttons to be available:

1. Something must be down the side of the table. You must put a variable in the Row list unless you specify that statistics labels appear down the side.
2. For multi-page tables, something must be across the top, too. If you put a variable into the Layer list, you must also put one into the Column list, unless you specify that statistics labels appear across the top.

When a variable is highlighted in the Rows, Columns, or Layers box, you can use the Selected Variable controls at the right to modify its function in the table:

Defines cells identifies a grouping variable, each of whose categories forms a row or column in the table, or a separate subtable. You can display counts or various percentages for the categories of a grouping variable.

Is summarized identifies a summary variable, one for which summary statistics appear in the table. The first summary variable you specify defines the summary dimension. The default summary statistic is the mean and is not labeled; you can label it by selecting the summary variable and visiting the Edit Statistics subdialog.

While a summary dimension is specified, the words Summary Dimension appear above the variable list for that dimension, and you cannot select Is summarized for a variable in another dimension.

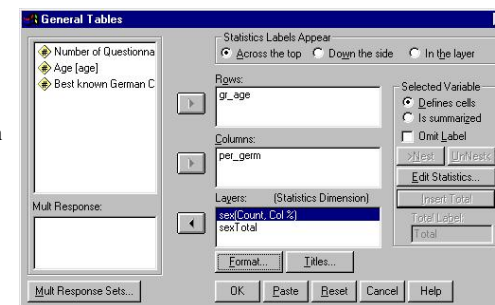


Figure 32: General Tables

Select [Omit Label] to suppress the variable label of the selected variable.

Multiple variables in any of these lists are stacked in their dimension within the table unless you nest them. The [Nest] button allows you to nest a selected variable within categories of the variable that precedes it in the list for its dimension. The [Unnest] button allows you to move a nested variable out of its nested position.

Nesting A nested variable is displayed "within" each category of a grouping variable or a multiple-response set. Up to 4 levels of nesting are allowed. You cannot nest a variable "within" a summarized variable or a total, and you cannot nest a multiple response set within another multiple-response set.

To nest a variable within categories of a variable below it in the list, remove it from the list and add it again at the bottom. For subgroup summary statistics, nest the summarized variable within categories of the grouping variable.

[Edit Statistics] lets you specify which statistics will be displayed for the selected variable. You cannot assign statistics to a variable within which another variable is nested. Once you have assigned statistics to a variable, they become the default statistics for other variables of the same type in the same dimension. This makes it easy to assign the same statistics to several variables in the statistics dimension.

[Insert Total] allows you to add total statistics after the selected variable's group statistics. You can modify the label that will be used for the total statistics in the Total Label: field. The same statistics are normally displayed for a total as for the variable totaled; but by selecting the total and clicking [Edit Statistics], you can specify different statistics for the line or column containing the total statistic(s).

Select [Format] to set margins, page dimensions, borders, and the treatment of missing cells and data.

Select [Titles] to specify lines of text for the table header, footer, and top-left corner.

Example: The MC sample survey

We will apply the above mentioned Tables commands to the MC sample survey .

```
*****
* Syntax file       : ex07_3.sps
* Data File charged : quest05.sav
* Output File      : res_07_3.lst
* Graphics File    :
* Comments         : Syntax and Output Files are displayed
* Comments         : in one file
* Comments         : The different section are marked by
* Comments         : the respective headers
* Author           : CDG-FZ München/klaus röder
*****
```

```
GET
FILE='C:\Roder_K\Roder_K53\Quest05.sav'.
EXECUTE .
```

* General Tables.

```
TABLES
/FORMAT BLANK MISSING(' ')
/GBASE=CASES
/FTOTAL= $t000001 "Total"
/TABLE=gr_age BY per_germ > (STATISTICS) BY sex + $t000001
/STATISTICS
count( sex( F5.0 ))
cpct( sex( PCT5.1 ) 'Col %':per_germ sex ).
```

* Results

Tables

(only results for female respondents are displayed)

Sex female

Age group	Best known German Personality											
	Gerhard Schröder		Boris Becker		Claudia Schiffer		Franz Beckenbauer		Helmut Kohl		Steffi Graf	
	Count	Col %	Count	Col %	Count	Col %	Count	Col %	Count	Col %	Count	Col %
<= 10												
10< <=20	1	20.0%	1	50.0%	1	100.0%	1	20.0%				
20< <=30							2	40.0%				
30< <=40	2	40.0%					2	40.0%			1	100.0%
40< <=50			1	50.0%								
50< <=60												
60< <=70												
70<	2	40.0%										

5.32.3. Tables of Frequencies

produces tables of the frequency distributions of categorical variables. Side-by-side frequency tables are available for multiple variables that share the same categories and for subgroups of cases.

Menu: [Statistics/Custom Tables/Tables of Frequencies]
Command Language:
TABLES

Move at least one variable in the Frequencies for: list. Multiple variables should normally have the same category codes and value labels. (If different value labels are present for the same value, SPSS uses the label from the variable specified first in this list.) To obtain frequency tables for Subgroups, move one or more variables into the appropriate list:
In Each Table displays the subgroup frequencies side-by-side within a single table.
Separate Tables displays a separate frequency table for each subgroup.
If you move two or more variables into either Subgroups list, choose between All combinations and Each variable separately.

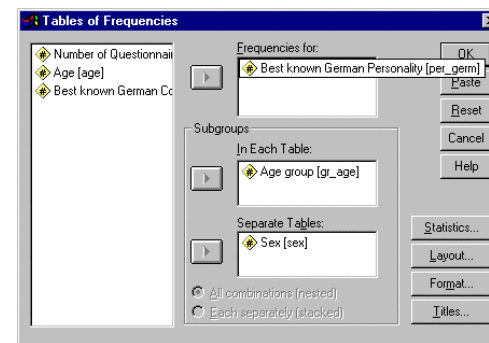


Figure 33: Tables of Frequencies

Select **[Statistics]** to choose the statistics for the frequency table
 Select **[Layout]** to indicate the layout of the frequency table.
 Select **[Format]** to set margins, page dimensions, borders, and the treatment of missing cells and data.
 Select **[Titles]** to specify lines of text for the table header, footer, and top-left corner.

Example: The MC sample survey

We will apply the above mentioned Tables commands to the MC sample survey .

```
*****
* Syntax file      : ex07_4.sps
* Data File charged : quest05.sav
* Output File     : res_07_4.lst
* Graphics File   :
* Comments        : Syntax and Output Files are displayed
* Comments        : in one file
* Comments        : The different section are marked by
* Comments        : the respective headers
* Author          : CDG-FZ München/klaus röder
*****
```

```
GET
FILE='C:\Roder_K\Roder_K53\Quest05.sav'.
EXECUTE .
```

```
* Table of Frequencies.
TABLES
/FORMAT BLANK MISSING('.') /TABLES
(LABELS) BY
gr_age > per_germ
/STATISTICS COUNT ((F5.0) 'Count' )
CPCT ((PCT7.1) '%' :gr_age ) .
```

* Results Tables

(results for the first 4 age groups in first table)

Age group	10< <=10	10< <=20	20< <=30	30< <=40
Gerhard Schröder				
Boris Becker				
Claudia Schiffer				
Franz Beckenbauer				
Helmut Kohl				
Steffi Graf				

	Best known German Personality Count	Best known German Personality % Count	Best known German Personality Count	Best known German Personality % Count
Gerhard Schröder	1	25.0%	1	14.3%
Boris Becker	1	25.0%		
Claudia Schiffer	1	25.0%		
Franz Beckenbauer	1	25.0%	2	28.6%
Helmut Kohl			2	28.6%
Steffi Graf	1	100.0%	2	28.6%

(results for the second 4 age groups in second table)

	40< <=50 Best known German Personality Count	50< <=60 Best known German Personality % Count	60< <=70 Best known German Personality Count	70< Best known German Personality Count
Gerhard Schröder				2
Boris Becker	3	75.0%		
Claudia Schiffer			2	66.7%
Franz Beckenbauer	1	25.0%		
Helmut Kohl		1 50.0%	1	33.3%
Steffi Graf		1 50.0%		

5.33. Graphics and Charts

Graphics are certainly the most attractive, comprehensive and persuasive means to communicate between the producer and consumer of statistical analysis. They are certainly recommended if little time or space is available to explain the thesis of the analyst. The publication of statistics via the media like television or newspapers is inconceivable without the presence of graphics. Graphs and charts however are usually not self-explained and the analyst must be careful to explain and comment the charts in an appropriate way, either by the possibilities of the graphs itself (legends etc.) or by simple text files. You should learn in this section the use of graphics, their modification and their importance in the context of the presentation of your data..

Creating Charts

There are mainly three ways of creating charts and plots in SPSS:

1. Commands from the Graphs menu on the main menu bar produce charts or plots. The main menu bar is available when the Data Editor, an output window, or a syntax window is active.
2. Certain statistical procedures produce plots. Examples include boxplots in the Explore procedure or scatterplots generated from Linear regression.
3. Interactive Charts

To create an interactive chart, select Interactive from the Graphs menu. For example, to create a bar chart, select Bar from the Interactive menu. Drag and drop variables from the source list to the target lists.

Modify Charts

There are mainly two ways of modifying charts and plots in SPSS:

1. Using the Chart Editor
2. Using the Interactive Charts E

5.33.1. Graphs from the Graphs Menu

To create a chart...

- Select [Graphs] from the menu bar.
- Select the type of chart you want from the Graphs menu.
- Choose the icon for the specific type of chart you want.
- You also need to indicate how your data are organized.

Options of Menu [Graphs]

- Bar allows you to generate a simple, clustered, or stacked bar chart from your data.
- Line allows you to generate a simple or multiple line chart from your data.
- Area allows you to generate a simple or stacked area chart from your data.
- Pie allows you to generate a simple pie chart or a composite bar chart from your data.
- High-Low allows you to plot pairs or triples of values, for example high, low, and closing prices.

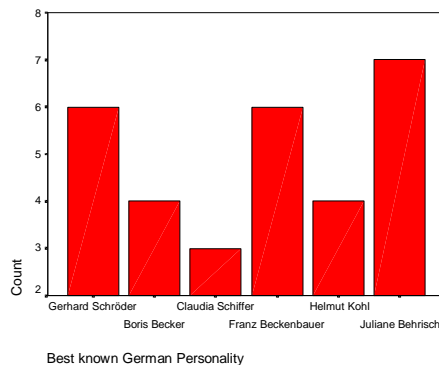


Figure 34: Graphs from the Graphs Menu

- Pareto creates Pareto charts, bar charts with a line superimposed showing the cumulative sum.
- Control produces the most commonly-used process-control charts.
- Boxplot allows you to generate boxplots showing the median, interquartile range, outliers, and extreme cases of individual variables.
- Error Bar allows you to generate boxplots showing the median, interquartile range, outliers, and extreme cases of individual variables.

- Scatter allows you to generate a simple or overlay scatterplot, a scatterplot matrix, or a 3-D scatterplot from your data.
- Histogram allows you to generate a histogram showing the distribution of an individual variable.
- Normal P-P plots the cumulative proportions of a variable's distribution against the cumulative proportions of the normal distribution.
- Normal Q-Q plots the quantiles of a variable's distribution against the quantiles of the normal distribution.
- Sequence produces a plot of one or more variables by order in the file, suitable for examining time-series data.
- Time Series: Autocorrelations calculates and plots the autocorrelation function (ACF) and partial autocorrelation function of one or more series to any specified number of lags, displaying the Box-Ljung statistic at each lag to test the overall hypothesis that the ACF is zero at all lags.
- Time Series: Cross-correlations calculates and plots the cross-correlation function of two or more series for positive, negative, and zero lags.

5.33.2. Graphs resulting from procedure calls

As in many previous examples, graphs can be produced as a result of SPSS procedures. Examples are the bar charts and histograms of the *FREQUENCIES* procedure.

5.33.3. Graphs from the Interactive Graphs Menu

A limited set of chart types can be selected through the Menu option: [Graphs/Interactive ..].

The following chart types are also available as interactive charts.

- Bar
- Line
- Pie
- Histogram
- Boxplot
- Error Bar
- Scatterplot

5.33.4. Using the Chart Editor

The Chart Editor is where you modify, print, and save charts that you have created elsewhere. You can change almost any attribute of a chart here. You can also change a chart from one type to another (for example, change a bar chart to a line chart).

To edit the overall chart, rather than a specific chart object, select a chart editor procedure from the Gallery, Chart, or Series menu.

To edit a chart object (such as an axis, title, legend item, bar or line), first select that object by clicking on it once. "Handles" appear on the selected object. (Instead of handles, a rectangle appears around selected text.) Selecting an object, such as a bar, selects all objects of that class. For example, selecting a bar in a simple bar chart selects all bars in that chart. Selecting one bar in a clustered bar chart selects all bars in that series.

After selecting the object, choose the appropriate palette for editing the attribute that you want to change. Select the palette in either of two ways:

1. Make a selection from the Attributes menu.
2. Select one of the attribute buttons on the chart window icon bar.

You can also select chart objects by double-clicking on them. When you double-click an object, an appropriate palette is automatically opened.

Attributes in Charts

All of the commands on the Attributes menu are available on the Icon bar when you are in the Chart Editor.

- Point selection. Click on the point selection tool to turn point selection mode on and off in the Chart window when a scatterplot or a boxplot is current.
- Fill Pattern allows you to change the fill of objects in an existing chart.
- Color allows you to change the color of a selected object.
- Markers allows you to change the markers on a chart or plot.
- Line Style allows you to change line styles.
- Bar Style allows you to change bar styles for charts that contain bars.
- Bar Label Style allows you to label bars with the numeric values that they represent.
- Interpolation allows you to select an interpolation method for a line or scatterplot series.
- Text allows you to change the font and point size of text within a chart or plot.
- Rotation rotates 3-D scatterplots, using a dialog box that shows only the axes.
- Swap Axes interchanges the vertical and horizontal axes of a chart.
- Explode Slice visually removes or "explodes" the selected slice from a pie chart.
- Break Line at Missing breaks a line chart where missing values should be.
- Chart options. Allows you to modify an existing chart.

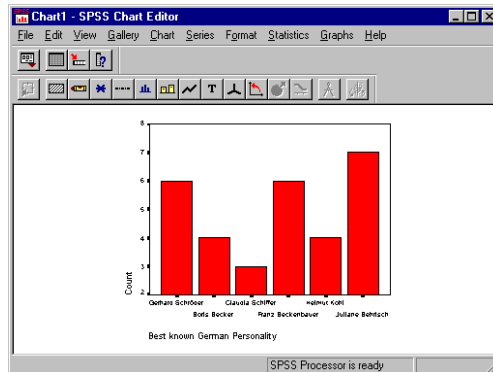


Figure 35: The Chart Editor

- Spin Mode rotates 3-D scatterplots directly, hiding only the labels. It is slower than 3-D Rotation when many data points are present.

5.33.5. Interactive Graphs and the Interactive Chart Editor

To Select and Modify a Chart Object

Activate the chart (double-click it).

Right-click the object to be modified.

The object is selected and a context menu is displayed. You can choose an action or a dialog box from the context menu.

or

Click the object to be modified.

Select a command from the

menus at the top of the Viewer.

or

Double-click the object to be modified.

This displays a dialog box for editing the object you clicked. Double-clicking a graphical element displays the element dialog box. For example, double-clicking a bar displays the Bars dialog box.

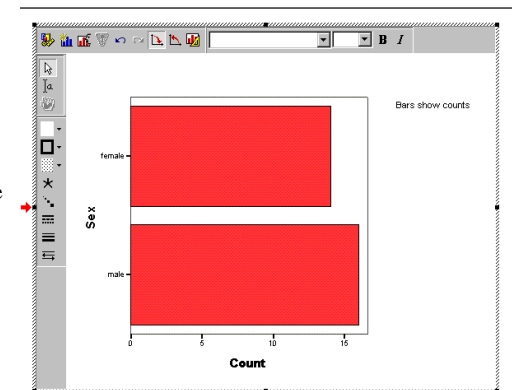


Figure 36: The Interactive Chart Editor

An interactive chart has access to the variables used to create it. Using the Assign Graph Variables dialog box, you can change variable assignments whenever the chart is activated.

If the data file that was used to create the chart is still open and nothing has changed in the Data Editor since the chart was created, the chart is labeled Interactive in the status bar of the Viewer, and you can change any of the variable assignments, using all of the variables available in the active data file. These charts use the Interactive Chart Editor.

If Save all data with the chart is selected on the Interactive tab of Options and the data file has been changed or a new data file has been opened, the chart is detached from the data file and is labeled Interactive (detached) in the status bar of the Viewer. For a detached chart, you can still change variable assignments; however, only variables used while the chart was actively connected to the open data file are now available. The Assign Graph Variables dialog box contains a message about available variables.

All interactive charts become static charts if Save only summarized data is selected on the Interactive tab of Options (Edit menu of the Viewer). In a Static chart, only the summarized

data is saved, and you cannot change variable assignments, although you can change attributes such as color, fill, and symbol size. You can turn an Interactive chart into a Static chart by right-clicking and selecting Save only summarized data from the context menu. A static chart cannot be changed to an interactive chart.

Non-interactive charts and charts created in previous versions of the software are not interactive. The chart objects can be edited for attributes such as color, but the variable assignments cannot be changed. These charts use the Chart Editor window available in previous releases of the program.

Exercise: The Cameroon Household Survey:

In the annex you will find a description of the Cameroon Household Survey. Use the data file CHS96_04.SAV

Use SPSS to present to following results:

Listing and Reports of:

1. Highest School Degree obtained and Age Groups
2. Medical Expenses by Age Group and Sex

Tables of:

3. Main Occupation of Head of Household (Frequency)
4. Main Occupation of Head of Household by Age Group and Sex
5. Main Occupation of Head of Household by Highest School Degree obtained

Charts of:

6. Reasons for not Attending School
7. Highest School Degree obtained by Main Occupation of Head of Household and Sex (3-D plot)
8. Employment Status and Medical Expenses