

Structure and design of user-friendly tables and graphs

Jan Erik Kristiansen

Senior adviser

Statistics Norway

Tables: Some general principles

Total population in EU-countries (version 1)			
	1994	1995	1996
B	10101123	10131231	10143325
DK	5197125	5216158	5251245
DK	81338215	81539256	81818261
GR	10410125	10443325	10465251
E	39121215	39177321	39242322
F	57779253	58020326	58256323
IRL	3580126	3595211	3616353
I	57139213	57269328	57333234
L	401122	407321	413256
NL	15432223	15424365	15494258
A	8015254	8040265	8055369
P	9892278	9912456	9921421
FIN	5078541	5099357	5117247
S	8745321	8816189	8838337
UK	58293211	58500293	58694362
Total	370524345	371592402	372660564

Version 2: Numbers right justified

	1994	1995	1996
B	10101123	10131231	10143325
DK	5197125	5216158	5251245
DK	81338215	81539256	81818261
GR	10410125	10443325	10465251
E	39121215	39177321	39242322
F	57779253	58020326	58256323
IRL	3580126	3595211	3616353
I	57139213	57269328	57333234
L	401122	407321	413256
NL	15432223	15424365	15494258
A	8015254	8040265	8055369
P	9892278	9912456	9921421
FIN	5078541	5099357	5117247
S	8745321	8816189	8838337
UK	58293211	58500293	58694362
Total	370524345	371592402	372660564

Version 3: Grouping digits

	1994	1995	1996
B	10 101 123	10 131 231	10 143 325
DK	5 197 125	5 216 158	5 251 245
D	81 338 215	81 539 256	81 818 261
GR	10 410 125	10 443 325	10 465 251
E	39 121 215	39 177 321	39 242 322
F	57 779 253	58 020 326	58 256 323
IRL	3 580 126	3 595 211	3 616 353
I	57 139 213	57 269 328	57 333 234
L	401 122	407 321	413 256
NL	15 432 223	15 424 365	15 494 258
A	8 015 254	8 040 265	8 055 369
P	9 892 278	9 912 456	9 921 421
FIN	5 078 541	5 099 357	5 117 247
S	8 745 321	8 816 189	8 838 337
UK	58 293 211	58 500 293	58 694 362
Total	370 524 345	371 592 402	372 660 564

Version 4: Thousands

	1994	1995	1996
B	10 101	10 131	10 143
DK	5 197	5 216	5 251
DK	81 338	81 539	81 818
GR	10 410	10 443	10 465
E	39 121	39 177	39 242
F	57 779	58 020	58 256
IRL	3 580	3 595	3 616
I	57 139	57 269	57 333
L	401	407	413
NL	15 432	15 424	15 494
A	8 015	8 040	8 055
P	9 892	9 912	9 921
FIN	5 078	5 099	5 117
S	8 745	8 816	8 838
UK	58 293	58 500	58 694
Total	370 521	371 588	372 656

Version 5: Millions (1 decimal)

	1994	1995	1996
B	10,1	10,1	10,1
DK	5,2	5,2	5,3
DK	81,3	81,5	81,8
GR	10,4	10,4	10,5
E	39,1	39,2	39,2
F	57,8	58,0	58,3
IRL	3,6	3,6	3,6
I	57,1	57,3	57,3
L	0,4	0,4	0,4
NL	15,4	15,4	15,5
A	8,0	8,0	8,1
P	9,9	9,9	9,9
FIN	5,1	5,1	5,1
S	8,7	8,8	8,8
UK	58,3	58,5	58,7
Total	370,5	371,6	372,7

Version 6: Dividers

	1994	1995	1996
B	10,1	10,1	10,1
DK	5,2	5,2	5,3
DK	81,3	81,5	81,8
GR	10,4	10,4	10,5
E	39,1	39,2	39,2
F	57,8	58,0	58,3
IRL	3,6	3,6	3,6
I	57,1	57,3	57,3
L	0,4	0,4	0,4
NL	15,4	15,4	15,5
A	8,0	8,0	8,1
P	9,9	9,9	9,9
FIN	5,1	5,1	5,1
S	8,7	8,8	8,8
UK	58,3	58,5	58,7
Total	370,5	371,6	372,7

Version 7: Bold

	1994	1995	1996
B	10,1	10,1	10,1
DK	5,2	5,2	5,3
DK	81,3	81,5	81,8
GR	10,4	10,4	10,5
E	39,1	39,2	39,2
F	57,8	58,0	58,3
IRL	3,6	3,6	3,6
I	57,1	57,3	57,3
L	0,4	0,4	0,4
NL	15,4	15,4	15,5
A	8,0	8,0	8,1
P	9,9	9,9	9,9
FIN	5,1	5,1	5,1
S	8,7	8,8	8,8
UK	58,3	58,5	58,7
Total	370,5	371,6	372,7

Version 8: Per cent added

	1994	1995	1996	1996 %
B	10,1	10,1	10,1	2,7
DK	5,2	5,2	5,3	1,4
D	81,3	81,5	81,8	22,0
GR	10,4	10,4	10,5	2,8
E	39,1	39,2	39,2	10,5
F	57,8	58,0	58,3	15,6
IRL	3,6	3,6	3,6	1,0
I	57,1	57,3	57,3	15,4
L	0,4	0,4	0,4	0,1
NL	15,4	15,4	15,5	4,2
A	8,0	8,0	8,1	2,2
P	9,9	9,9	9,9	2,7
FIN	5,1	5,1	5,1	1,4
S	8,7	8,8	8,8	2,4
UK	58,3	58,5	58,7	15,8
Total	370,5	371,6	372,7	100,0

Version 9: Sorted, decending order

	1994	1995	1996	1996 %
D	81,3	81,5	81,8	22,0
UK	58,3	58,5	58,7	15,8
F	57,8	58,0	58,3	15,6
I	57,1	57,3	57,3	15,4
E	39,1	39,2	39,2	10,5
NL	15,4	15,4	15,5	4,2
GR	10,4	10,4	10,5	2,8
B	10,1	10,1	10,1	2,7
P	9,9	9,9	9,9	2,7
S	8,7	8,8	8,8	2,4
A	8,0	8,0	8,1	2,2
DK	5,2	5,2	5,3	1,4
FIN	5,1	5,1	5,1	1,4
IRL	3,6	3,6	3,6	1,0
L	0,4	0,4	0,4	0,1
Total	370,5	371,6	372,7	100,0

Example:

Table 14: Livestock Population 1994 – 1999

	1994	1995	1996	1997	1998	1999
Poultry	56,373	86,165	107096	220559	101270	46281
Goats	2164	2876	5603	5781	6803	4113
Cattle	43528	49476	47941	47381	55936	52940
Pigs	6175	9074	3906	19318	19062	18142
Sheep	1593	1819	3906	1886	2063	586
Rabbits	207	1402	2511	488	505	97

From: Karonga District Socio Economic Profile

Better: Numbers justified to the right, text to the left

Table 14. Livestock population 1994-1999						
	1994	1995	1996	1997	1998	1999
Poultry	56,373	86,165	107,096	220,56	101,270	46,281
Goats	2,164	2,876	5,603	5,781	6,803	4,113
Cattle	43,528	49,476	47,941	47,381	55,936	52,940
Pigs	6,175	9,074	3,906	19,318	19,062	18,142
Sheep	1,593	1,819	3,906	1,886	2,063	586
Rabbits	207	1,402	2,511	488	505	97

Even better: Change columns and rows. It is now much easier to compare years!

Table 14. Livestock population 1994-1999						
	Poultry	Goats	Cattle	Pigs	Sheep	Rabbits
1994	56,373	2,164	43,528	6,175	1,593	207
1995	86,165	2,876	49,476	9,074	1,819	1,402
1996	107,096	5,603	47,941	3,906	3,906	2,511
1997	220,559	5,781	47,381	19,318	1,886	488
1998	101,270	6,803	55,936	19,062	2,063	505
1999	46,281	4,113	52,940	18,142	586	97

Variables

**Independent
variable**



**Dependent
variable**

Classification variable

Cause

Often “invariant”

Indicator

Effect

“What we measure”

Examples of...

- **Classification/independent/background variables:**
 - Sex
 - Age
 - Household size
 - Locality
 - Region
 - Schooling
 - Occupation
 - Labour force participation
 - etc.
- **Indicators/dependent variables:**
 - Labour force participation rate
 - CPI
 - Newspaper reading
 - Infant mortality rate
 - Educational level
 - Housing standard
 - Fertility
 - etc, etc.

Percentage tables

- Percentages should be done on basis of the independent (classification) variable and percentages should then be compared across this variable.

Example:

Watched television on an average day
Did not watch television

		Compare	
Males			Female
85			79
15			21
<hr/>		<hr/>	
100%			100%

Decimals when reporting percentages

- Never use two decimals when giving percentages
- When reporting percentages from censuses, registers, etc., use one decimal
- When reporting percentages from surveys, use no decimal (unless the sample is very big: LF survey)

Two effective digits

- **When reporting numbers in the text, rounding is often effective:**

“The number of female students increased from 28 972 to 94 354, while the number of male students rose from 45 362 to 89 136”

“The number of female students increased from 29 000 to 94 000, while the number of male students rose from 45 000 to 89 000”

Simplify titles (examples)

Not: Table 3.4. Percentage distribution of households by main type of household. 1997

But: 3.4. Household types. 1997. Per cent

Not: Percentage distribution of persons aged 15-49 years with knowledge of transmission of Aid virus....

But: Persons 15-49 years with knowledge of.....

Or: Knowledge of Persons aged 15-49.

Simplify titles (more examples)

Not: Percentage distribution of persons outside the labour force by sex, age and reason

But: Persons outside the labour force by reason, sex and age. Per cent

Not: Pupil to teacher ratio

But: Pupils per teacher

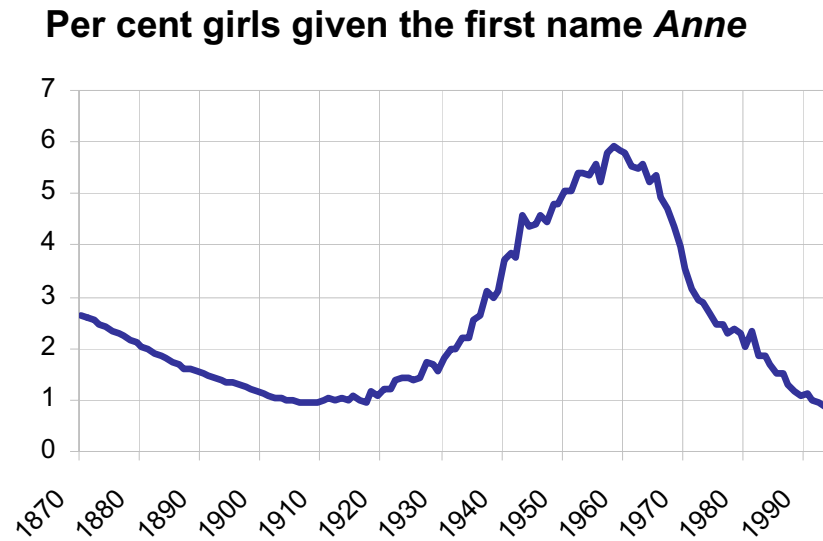
Visual statistics

**From tables to graphs:
Why and how?**

Example:

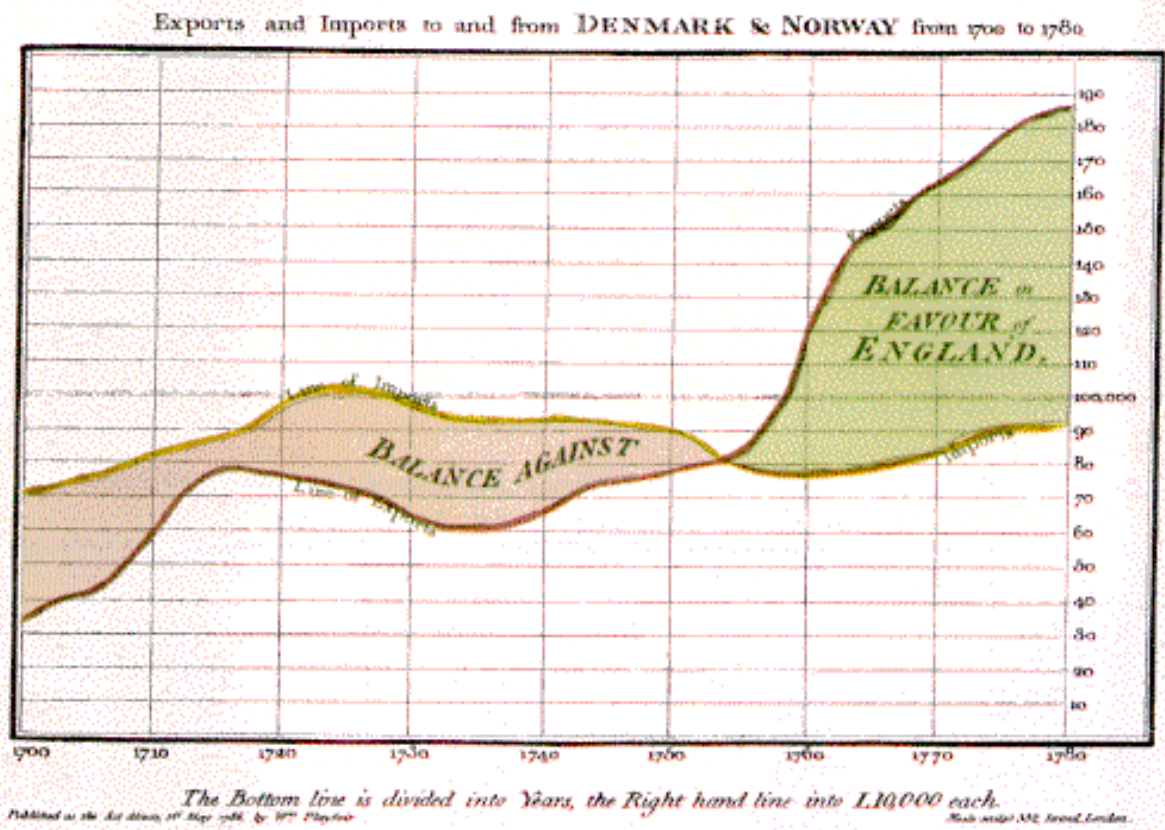
Per cent girls given the first name *Anne* each year

Year:	Pct.:	Year:	Pct.:	Year:	Pct.:	Year:	Pct.:	Year:	Pct.:	Year:	Pct.:
1870	2,65	1891	1,47	1912	0,99	1933	2,22	1954	5,35	1975	2,48
1871	2,59	1892	1,43	1913	1,03	1934	2,21	1955	5,57	1976	2,45
1872	2,53	1893	1,39	1914	0,98	1935	2,54	1956	5,25	1977	2,27
1873	2,47	1894	1,35	1915	1,09	1936	2,65	1957	5,80	1978	2,38
1874	2,41	1895	1,32	1916	1,00	1937	3,13	1958	5,91	1979	2,30
1875	2,35	1896	1,28	1917	0,95	1938	2,97	1959	5,83	1980	2,02
1876	2,29	1897	1,24	1918	1,17	1939	3,12	1960	5,81	1981	2,33
1877	2,23	1898	1,20	1919	1,09	1940	3,70	1961	5,55	1982	1,87
1878	2,17	1899	1,16	1920	1,23	1941	3,85	1962	5,50	1983	1,87
1879	2,11	1900	1,13	1921	1,22	1942	3,75	1963	5,58	1984	1,69
1880	2,05	1901	1,09	1922	1,39	1943	4,59	1964	5,22	1985	1,51
1881	1,98	1902	1,05	1923	1,42	1944	4,35	1965	5,37	1986	1,51
1882	1,92	1903	1,03	1924	1,41	1945	4,39	1966	4,94	1987	1,30
1883	1,86	1904	1,01	1925	1,39	1946	4,59	1967	4,73	1988	1,18
1884	1,80	1905	0,98	1926	1,41	1947	4,46	1968	4,36	1989	1,09
1885	1,74	1906	0,96	1927	1,74	1948	4,78	1969	3,99	1990	1,14
1886	1,68	1907	0,94	1928	1,68	1949	4,80	1970	3,53	1991	0,99
1887	1,62	1908	0,94	1929	1,57	1950	5,07	1971	3,16	1992	0,97
1888	1,58	1909	0,94	1930	1,80	1951	5,05	1972	2,92	1993	0,88
1889	1,54	1910	0,98	1931	2,00	1952	5,38	1973	2,88		
1890	1,51	1911	1,02	1932	1,98	1953	5,42	1974	2,66		



Charts speak directly to the eye!

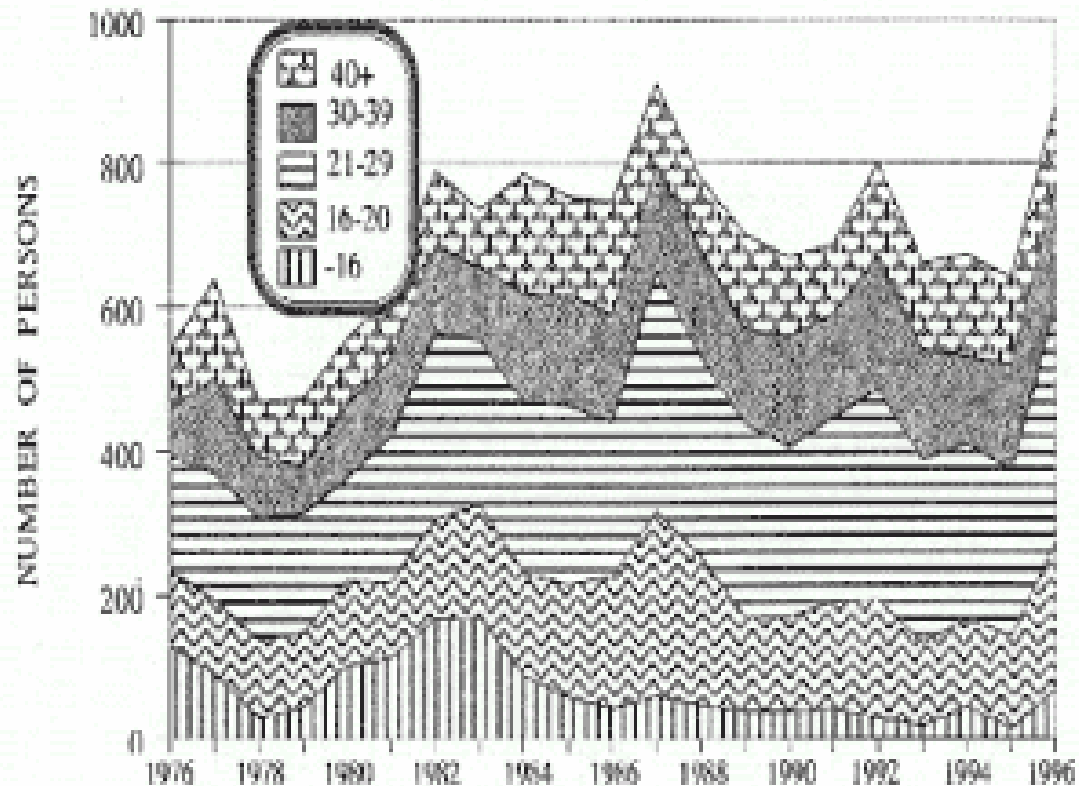
Example of good graphs:



William Playfair
1786!

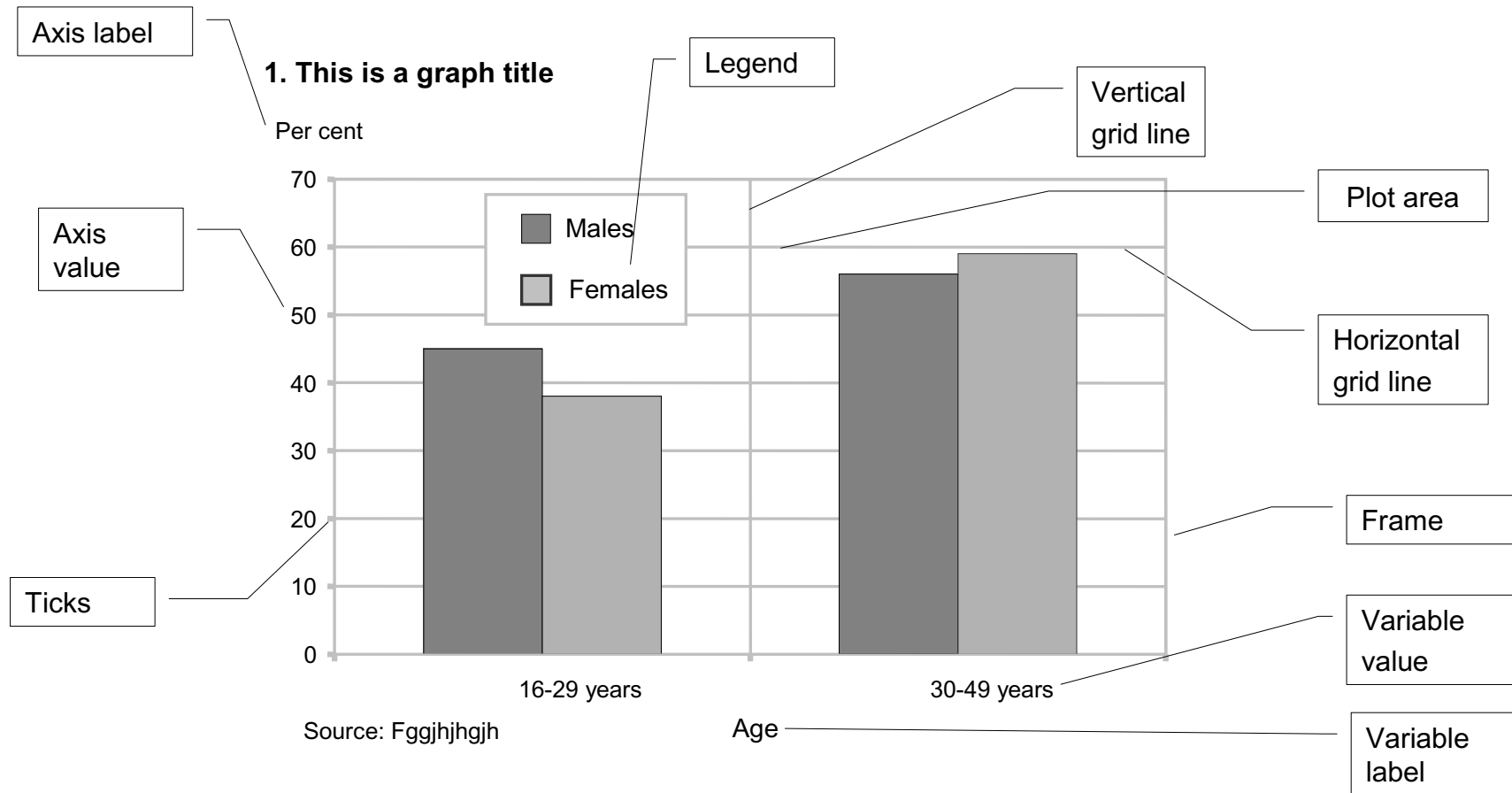
Example of poor graphs:

FIGURE 15: JUVENILES AND ADULTS CONVICTED BY AGE
(SERIOUS OFFENCES) 1976-1996



Cyprus Statistical
Abstract 1996

Elements of a graph:

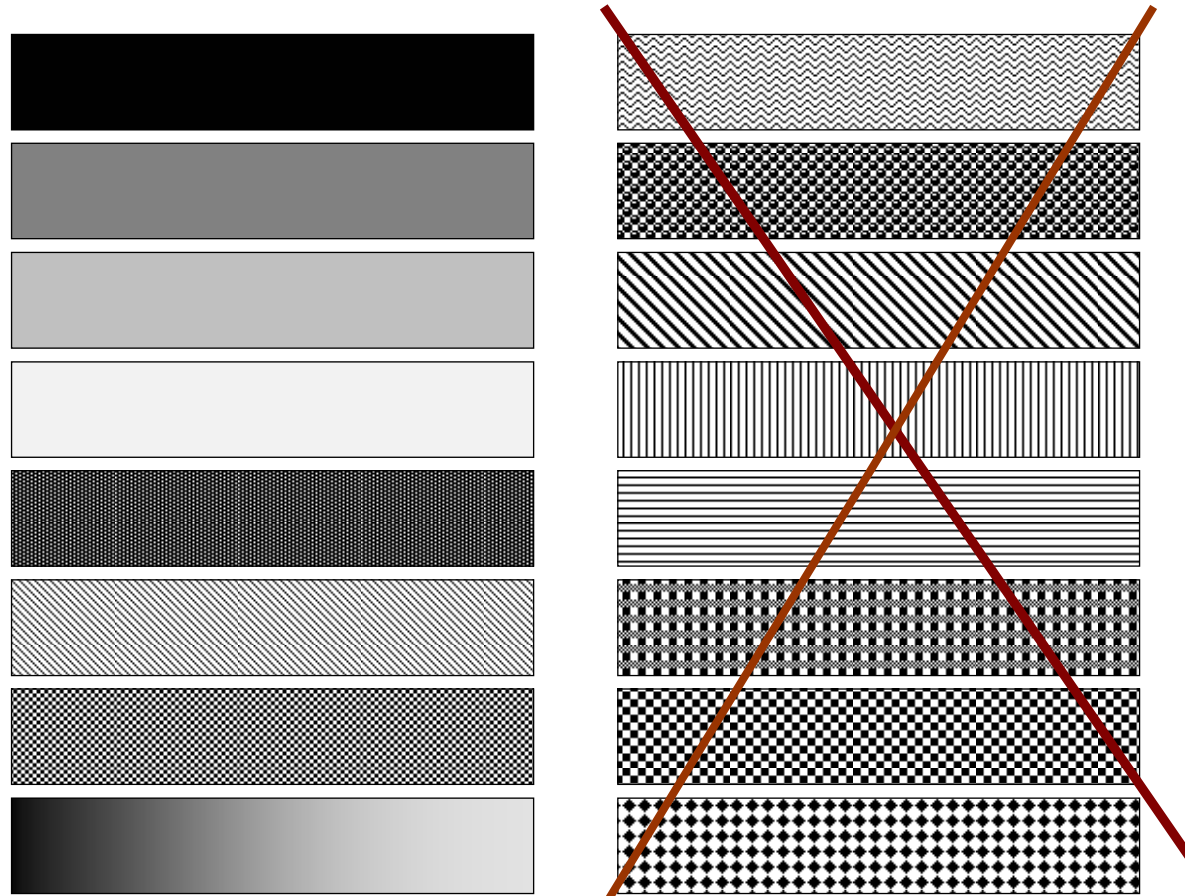


Graphs: Issues and solutions

- General recommendations:
 - Keep It Simple and Small (KISS)
 - Title: 10 points (bold)/ other text: 8 points normal
 - Outer frames not necessary
 - If possible, bring legend into graph (“plot area”)
 - Don’t use fancy patterns
 - Not too many values, groups, series (max. 5?)
- But: Solutions are dependent on the graphic program used!

Graphs: Issues and solutions (cont.)

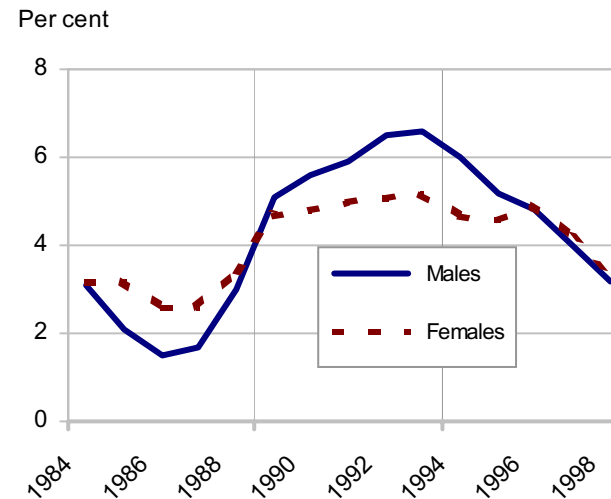
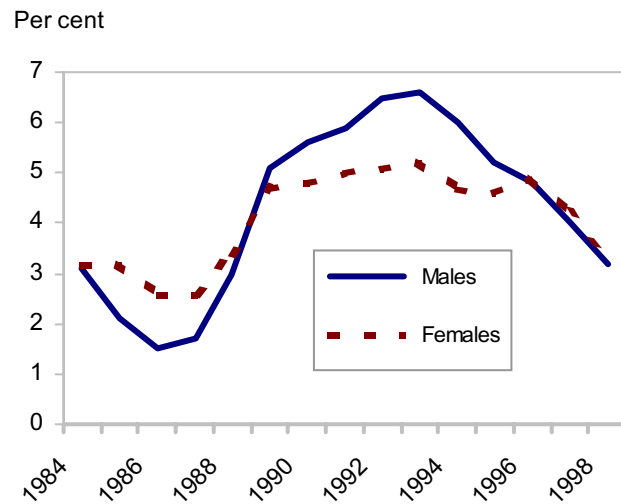
- Use of patterns



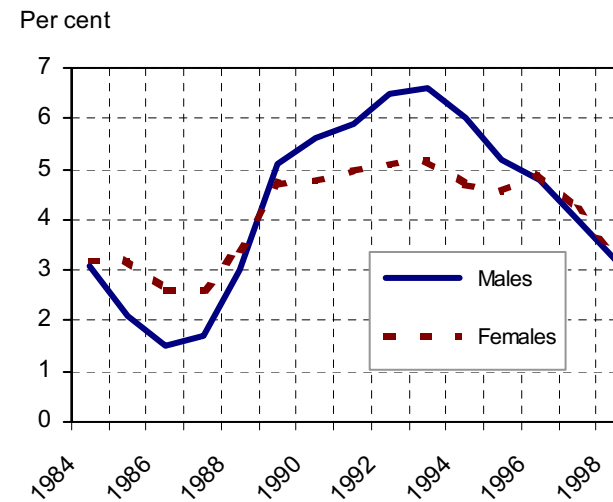
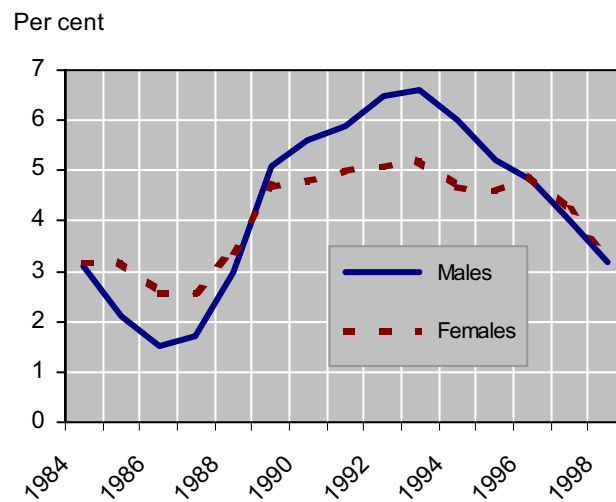
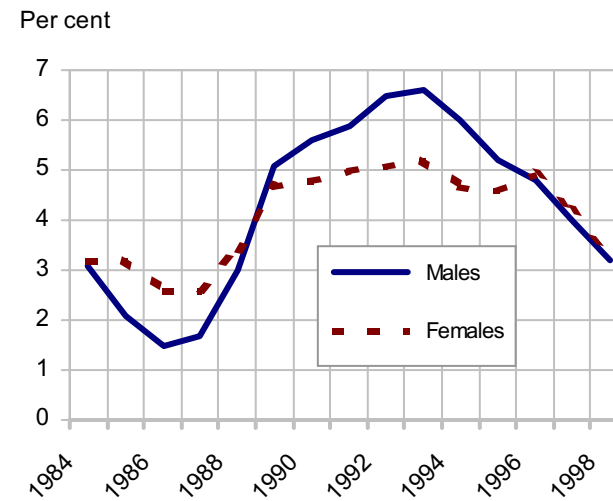
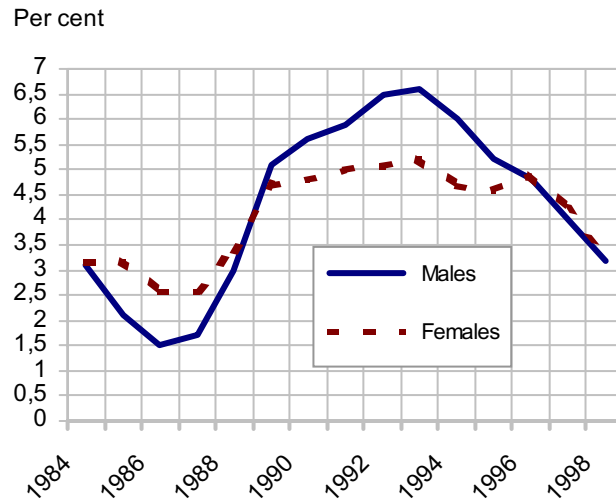
Maputo February 2005

Graphs: Issues and solutions (cont.)

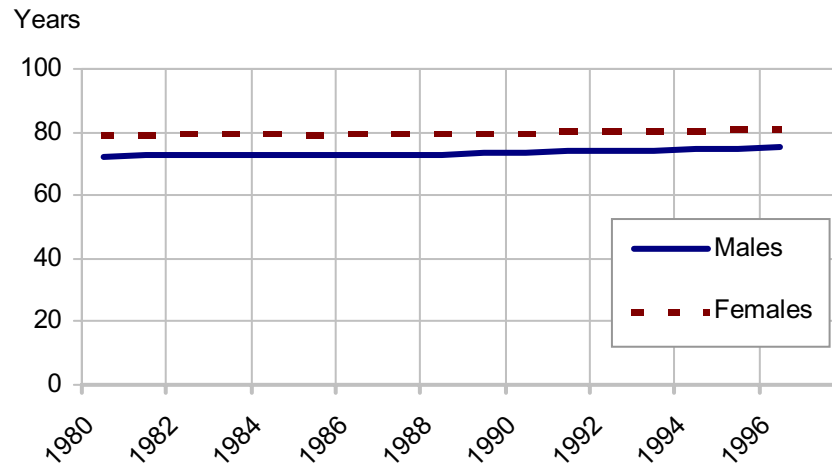
- Use gridlines to improve readability
- But how many?



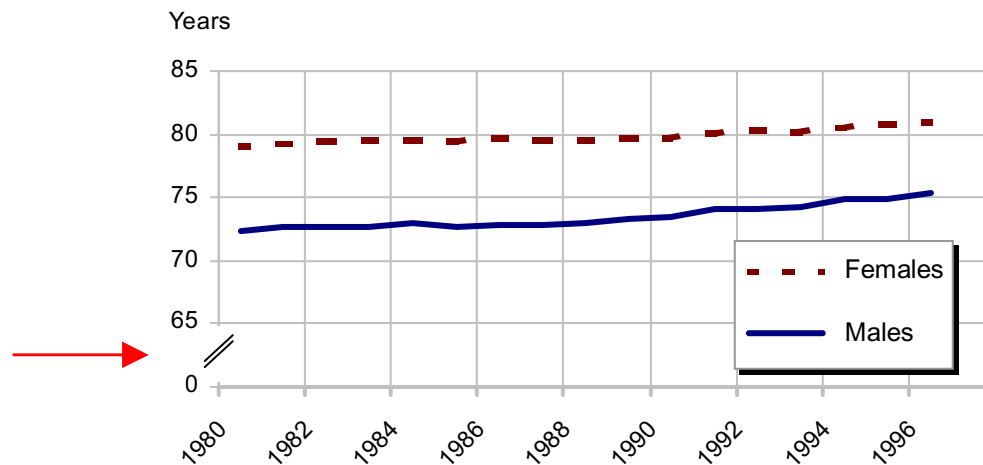
- Use of gridlines (cont.):



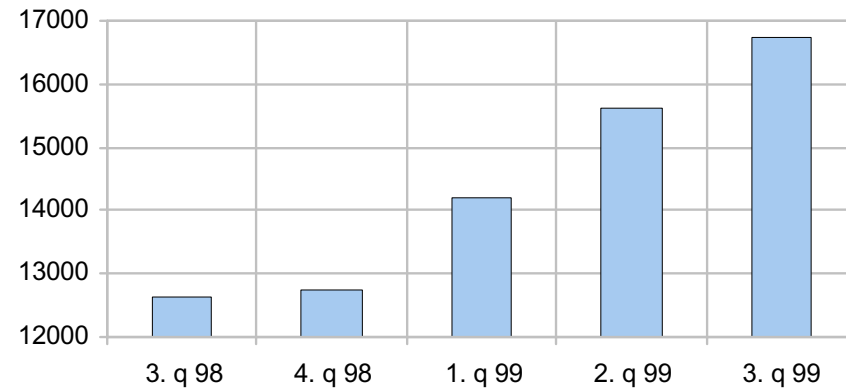
Graphs: Issues and solutions (cont.)



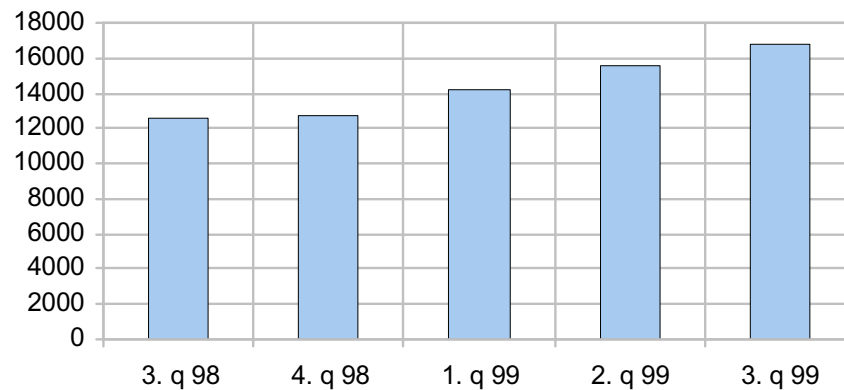
- 0 on the value axis?



Graphs: Issues and solutions (cont.)



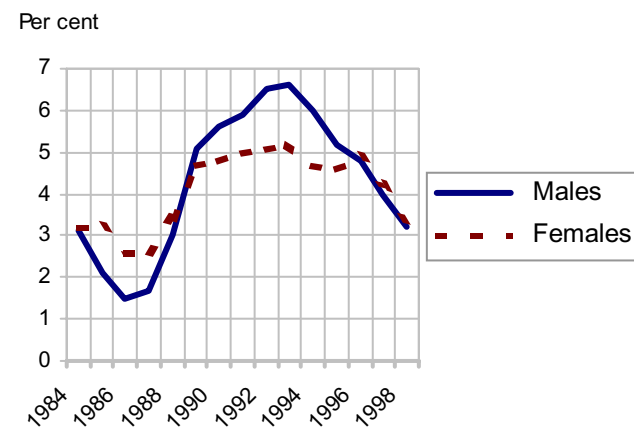
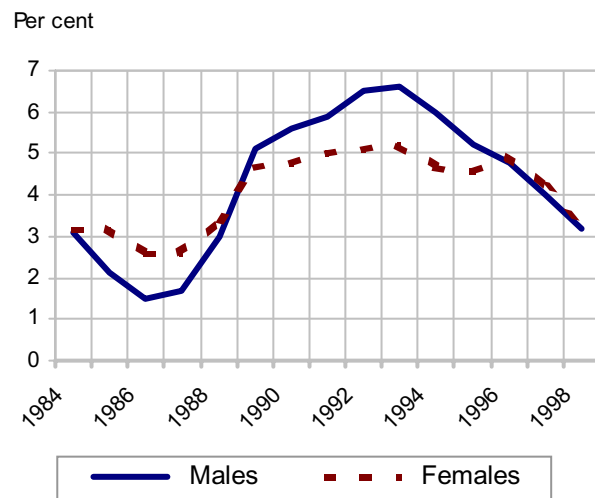
- Increase of 600 % ?!



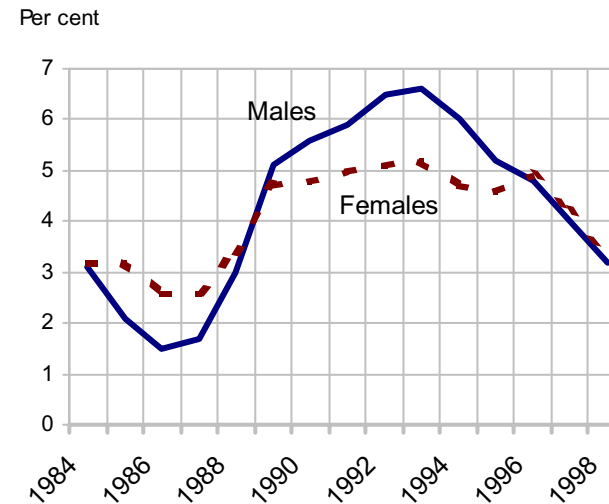
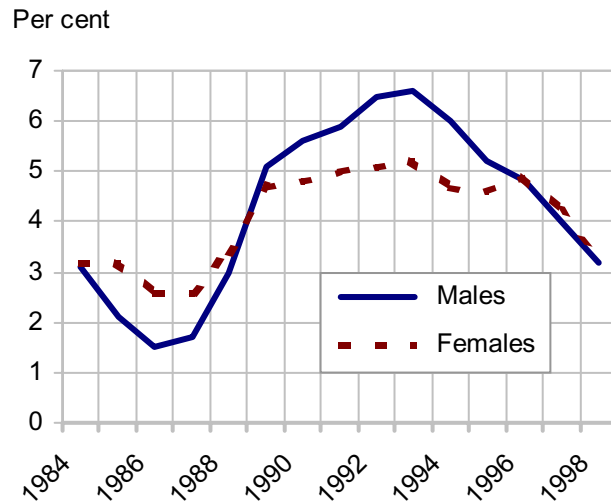
- No, only 35%!

Graphs: Issues and solutions (cont.)

- Legend: Where to place it?



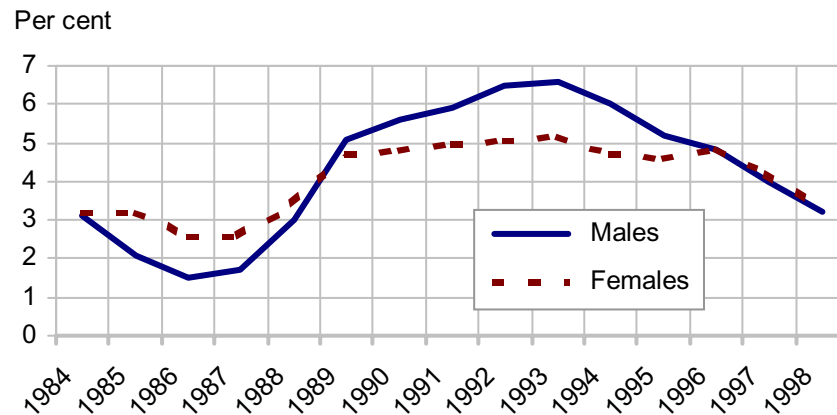
- Legend: Where to place it? (cont.)



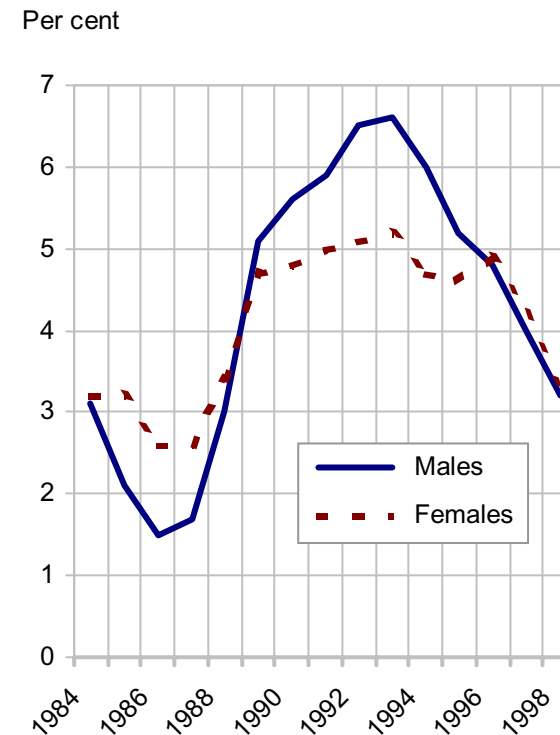
**Best
solution!**

Graphs: Issues and solutions (cont.)

- Format/proportions

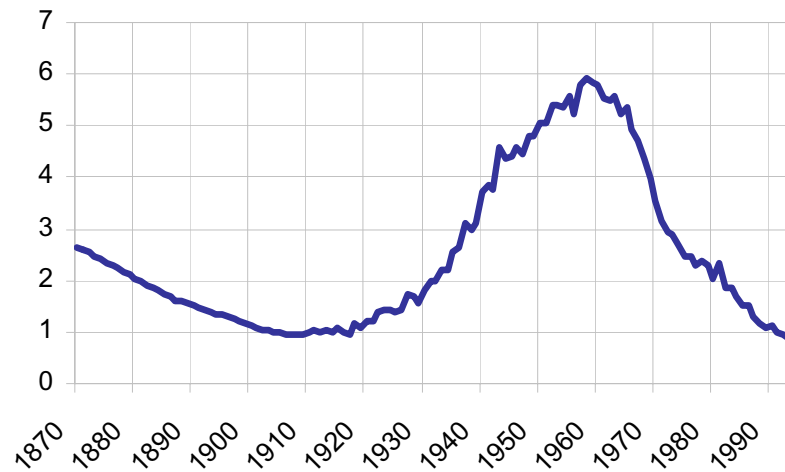


Landscape



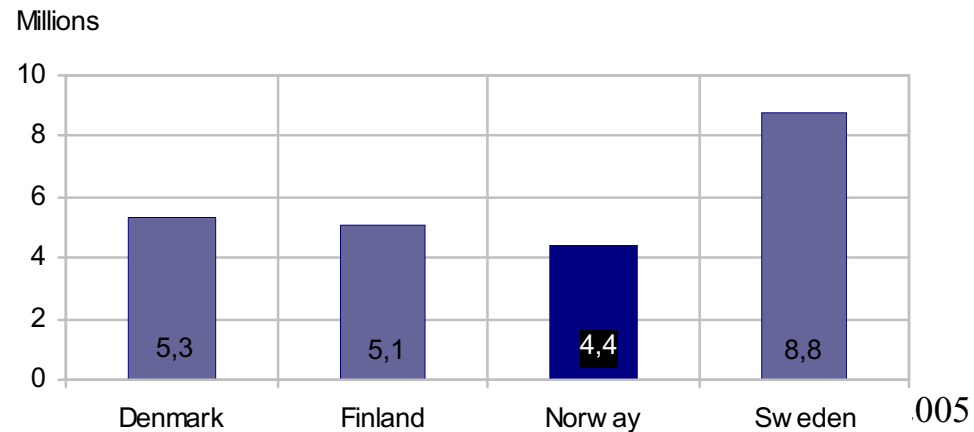
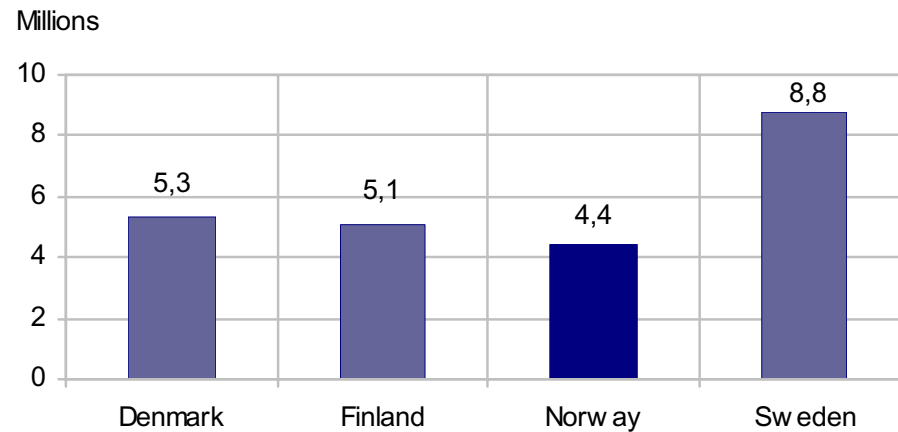
Portrait

- Format/proportions
 - The most usual format is probably landscape; width x height = ca. 1.5 x 1



Graphs: Issues and solutions (cont.)

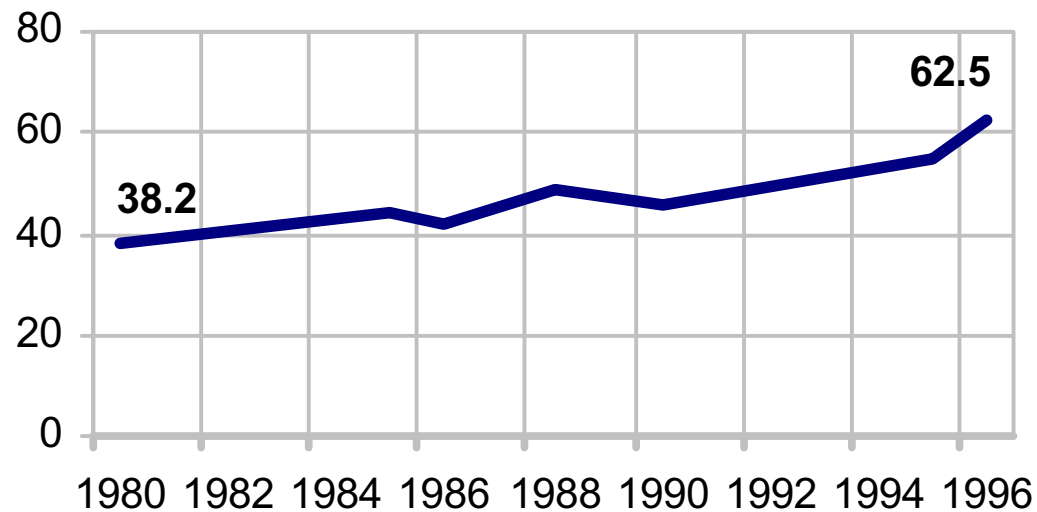
- Should values be given in the graph?



005

Graphs: Issues and solutions (cont.)

- In line charts, it is often useful to give the first and the latest value:



Main types of graphs:

- Bar charts (vertical)
 - Grouped
 - Stacked
- Bar charts (horizontal)
 - Grouped & stacked
- Line charts + area charts
- Pie charts
- Other types/ combinations

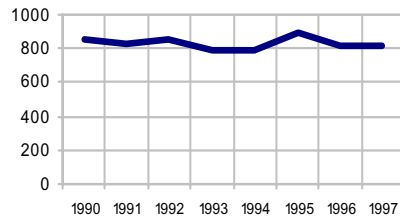
Which type of graph to use?:

Number of adopted children 1990-1997

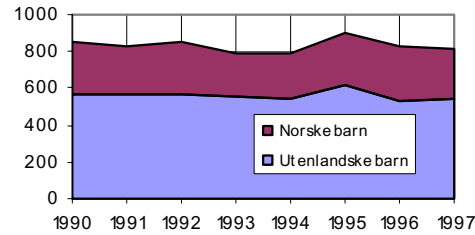
	Total				Norwegian			Foreign		
	Total	Under 3 years	3 - 11 years	12 years and over	Total	Boys	Girls	Total	Boys	Girls
1990	855	397	303	155	283	137	146	572	309	263
1991	833	418	263	152	271	131	140	562	309	253
1992	851	442	251	158	279	149	130	572	284	287
1993	786	397	246	143	236	125	111	550	312	238
1994	788	426	231	131	239	115	124	549	273	276
1995	898	490	257	151	284	141	143	614	272	342
1996	822	418	206	198	295	134	161	527	240	287
1997	814	469	189	156	272	126	146	542	244	298

Possible charts from the above table:

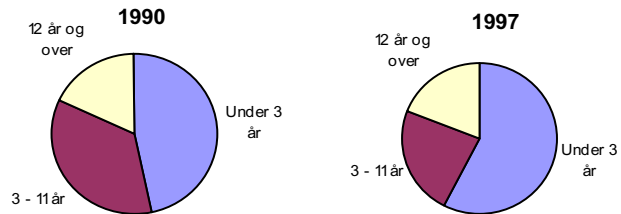
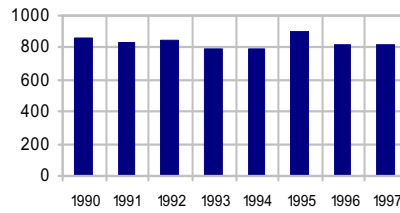
No. of adopted children 1990-1997



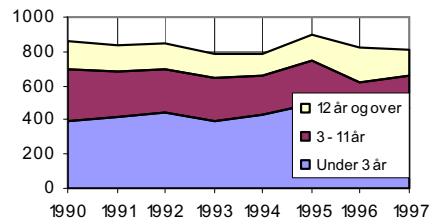
No. of adopted children. Norwegian and foreign 1990-1997



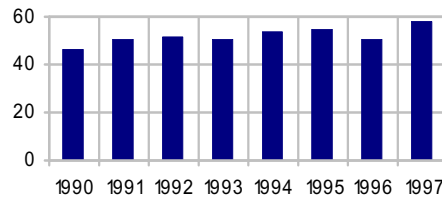
No. of adopted children 1990-1997



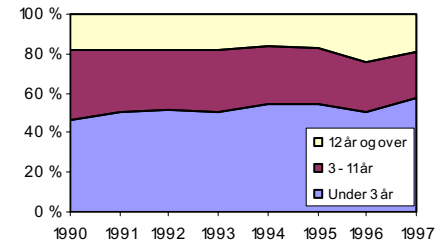
No. of adopted children by age 1990-1997



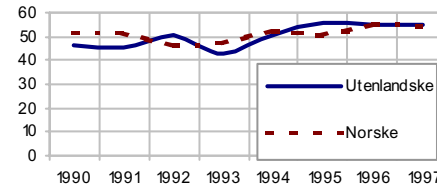
Adopted children. Percentage under the age of 3. 1990-1997



Adopted children by age 1990-1997

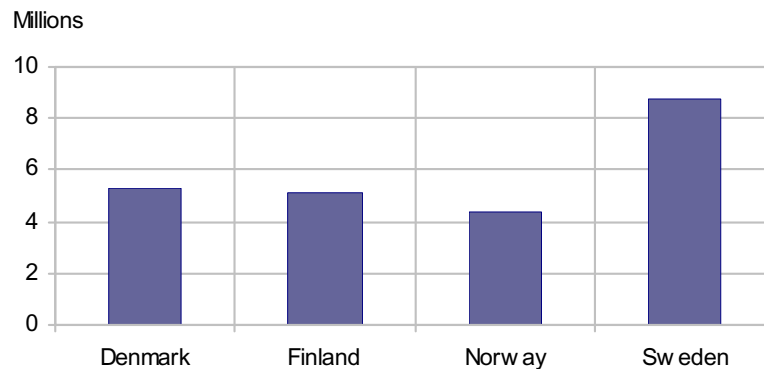


Adopted children. Percentage girls.

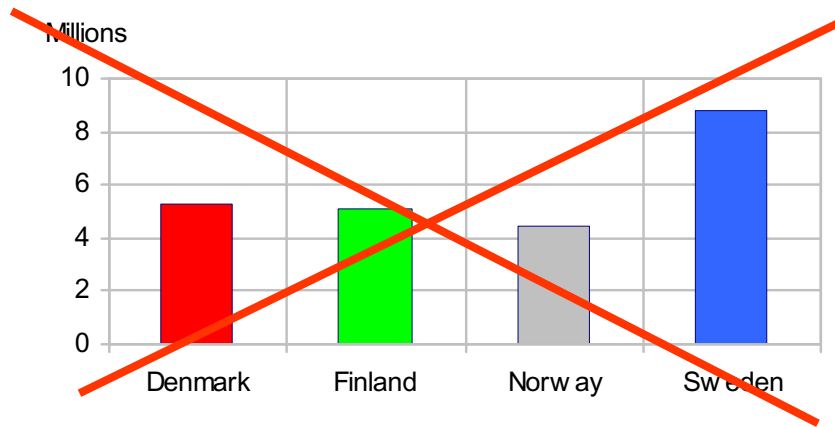


Bar charts (vertical):

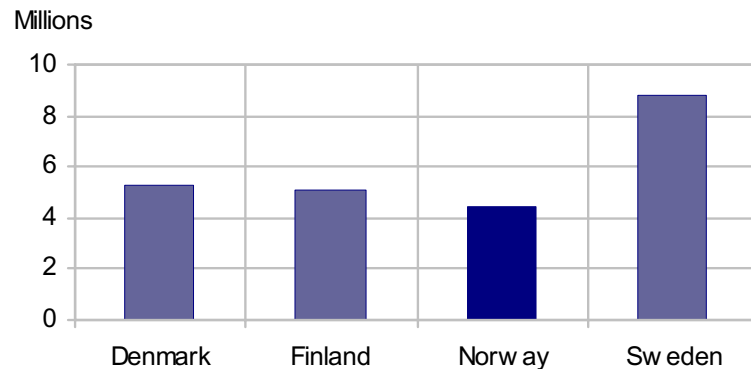
- Also called: Column charts
- The simplest and most basic chart type
- Used to show absolute or relative frequencies, percentages, averages
- Example: *Population in the Nordic countries*



Bar charts (vertical) (cont.):

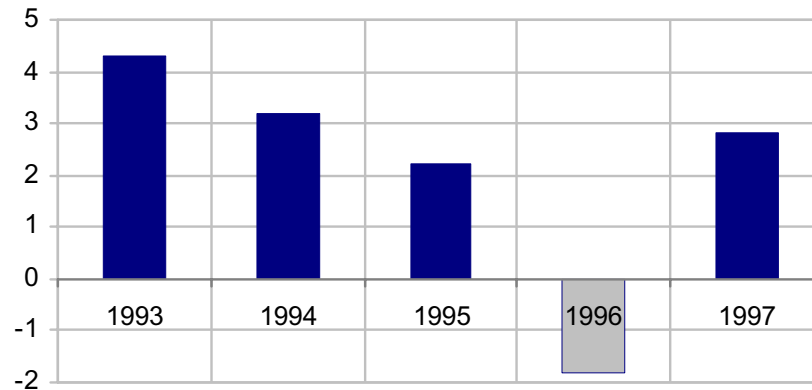


- With only one classification variable it is best to use only one colour/pattern



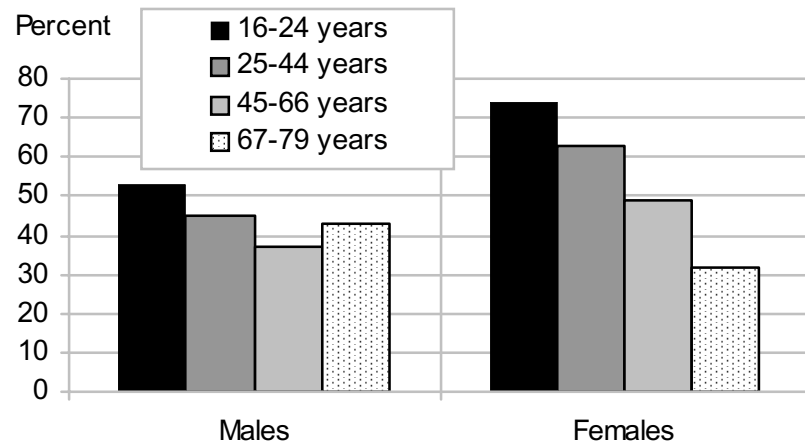
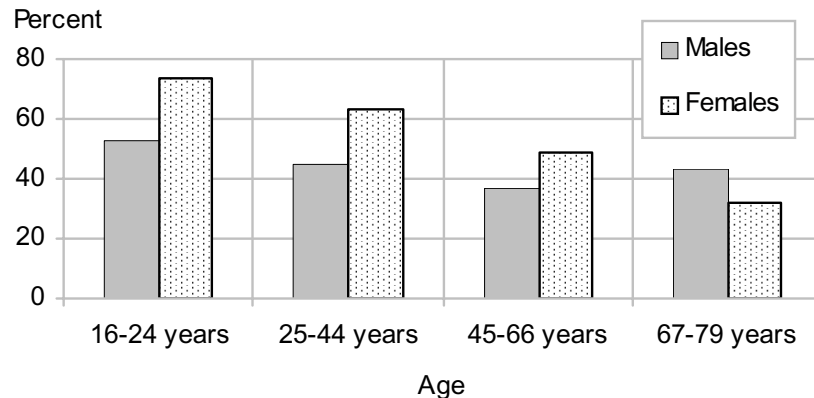
- If you want to focus on one class/group, this can be done by using a different colour/pattern

Bar charts (vertical) (cont.):



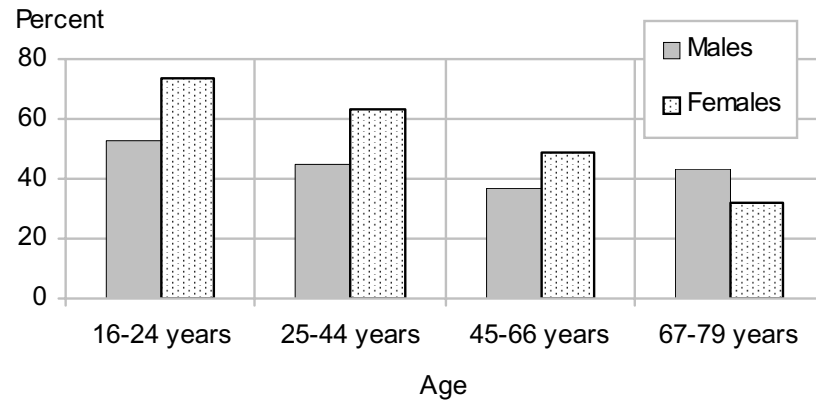
- When both positive and negative values are plotted, use different colours/ patterns

Grouped bar charts:



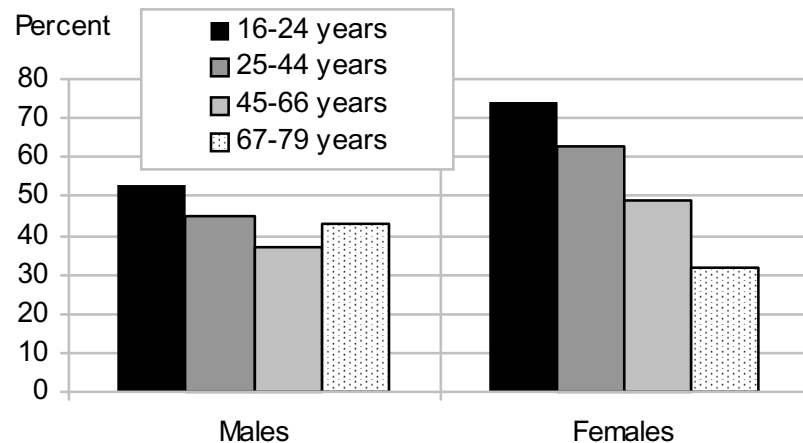
- Two (or more) categories
Example: *Percent using a library last year, by sex and age*
- With two categories, we have two ways of grouping, inviting to different comparison:

Grouped bar charts:



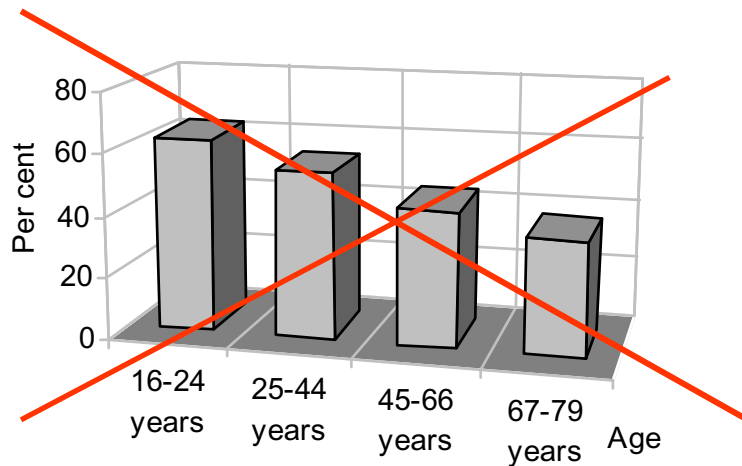
- This grouping invites a comparison of males and females in different age groups: “The difference between males and females are largest among the young”

Grouped bar charts:

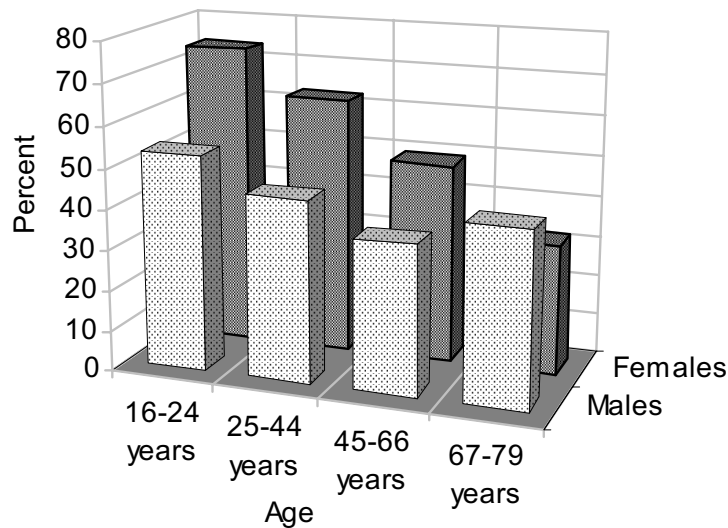


- This grouping invites a comparison of the effect of age, separately for the two sexes: “The percentage using a library decreases with age, but more so for females than males”.

“3D” bar charts:

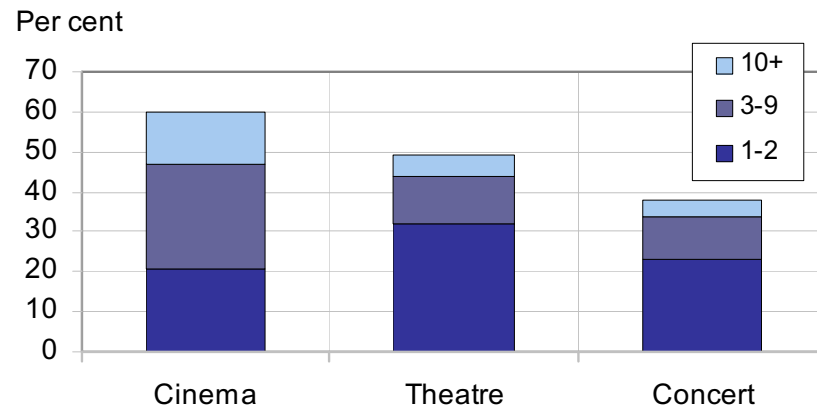


- Popular, but not recommended



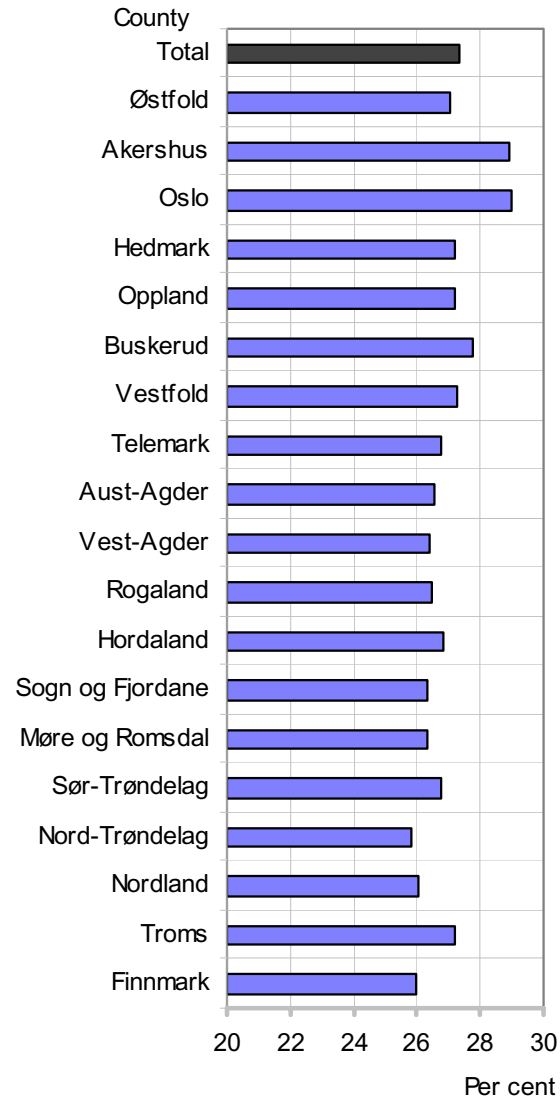
- “Genuine” 3D-chart (3 variables)

Stacked bar charts



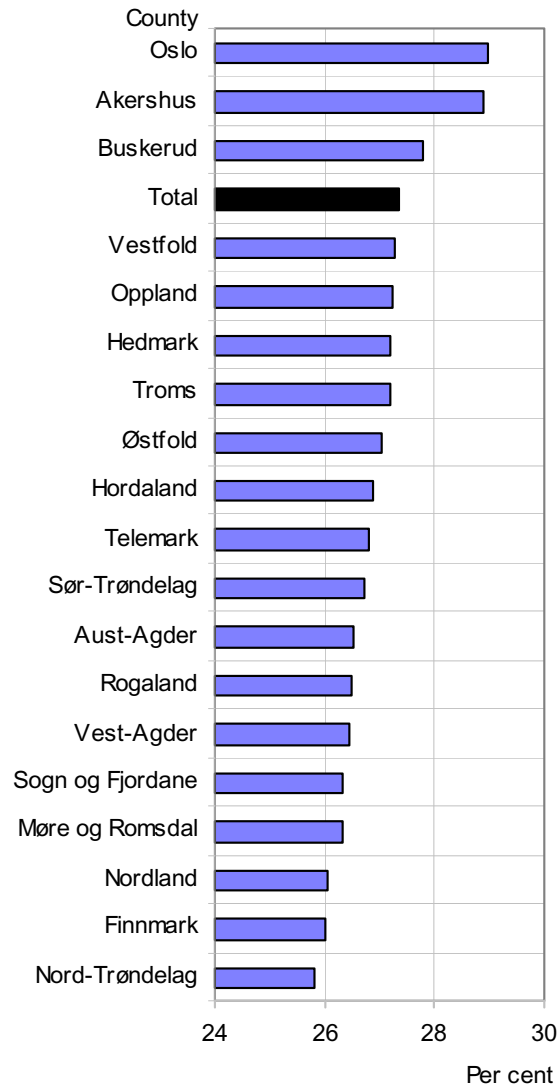
- Show total frequency and how the total is divided into different components

Horizontal bar charts:



- Often used
 - when labels (variable names) are long
 - when there are many variables or classifications
- Here, text should be right justified

Horizontal bar charts:

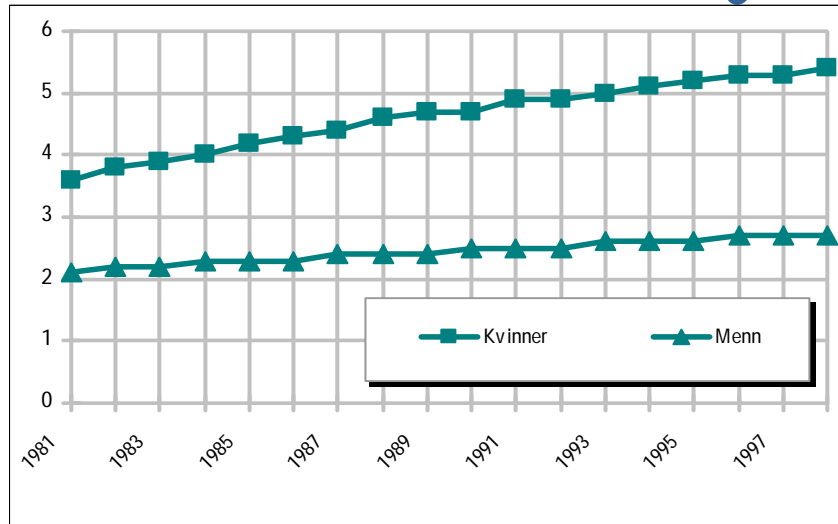


- Often it is better to sort by the value of the indicator (dependent variable)

Line charts:

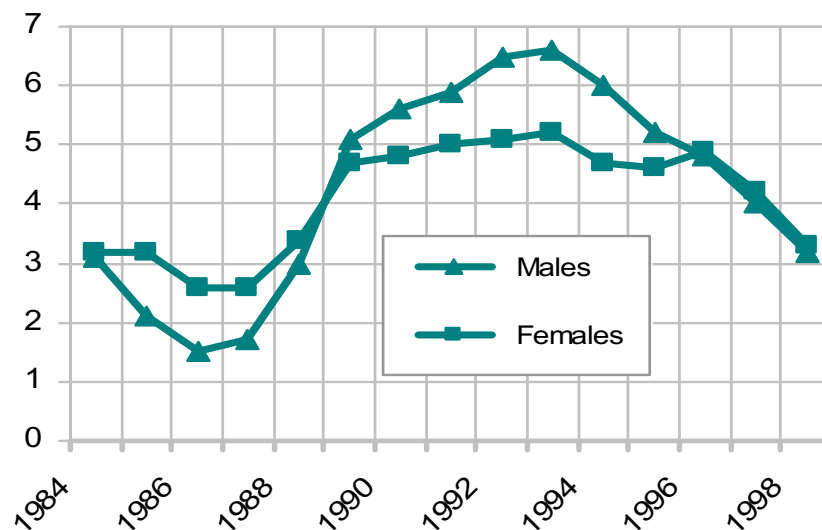
- Most often used for time series
- Time = years, quarters, months, weeks, days, hours and minutes + age
- What is a time “series”? Minimum = 4 data points?
Up to 4, use vertical bar chart
- The longer series, the better?!

Line charts: Symbols?



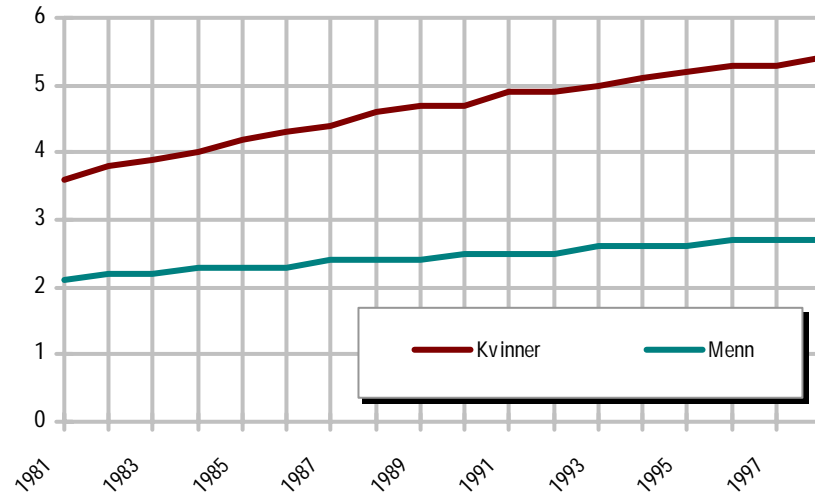
- Indicators or symbols (■ ▽) are often used to differentiate between series, but these symbols overload the chart, especially with crossing time series

Per cent

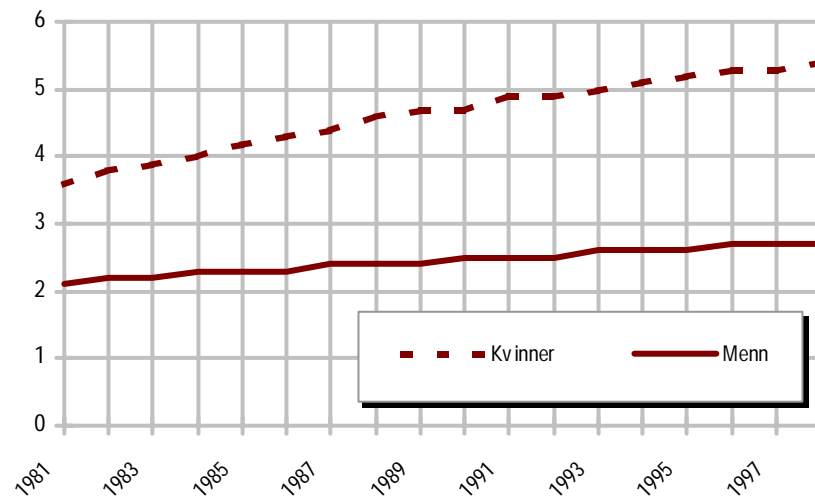


ary 2005

Line charts:

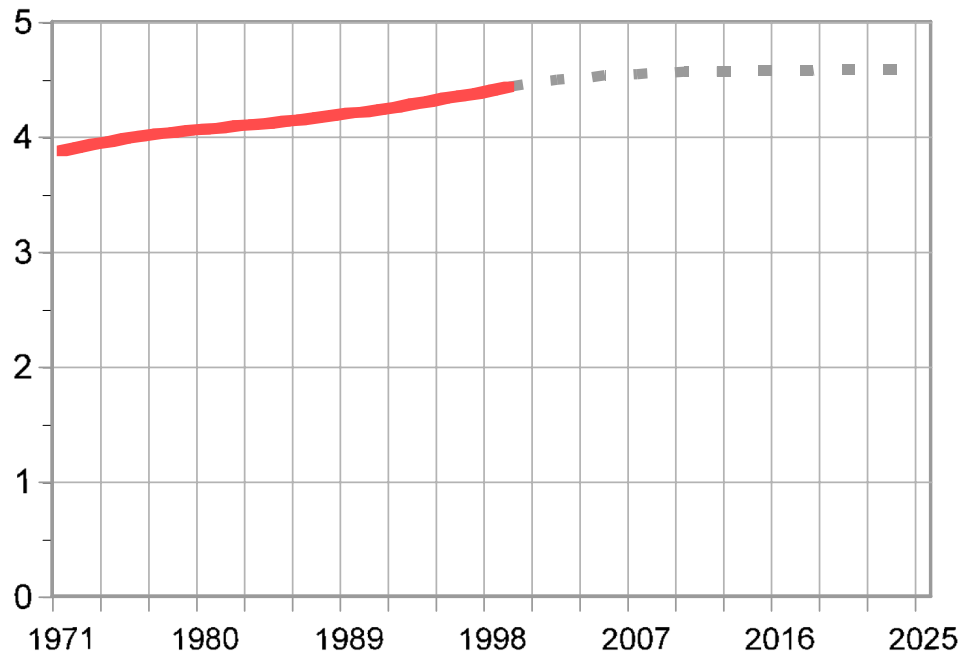


- Better to use different colours ...



- ... and/or line styles

Line charts:

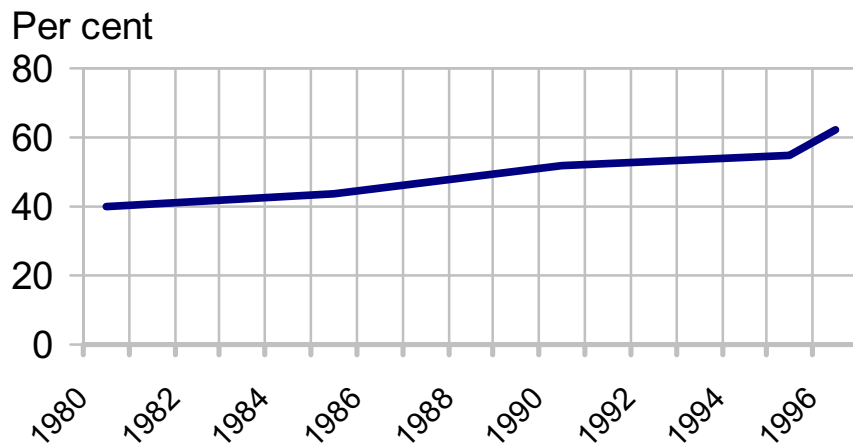
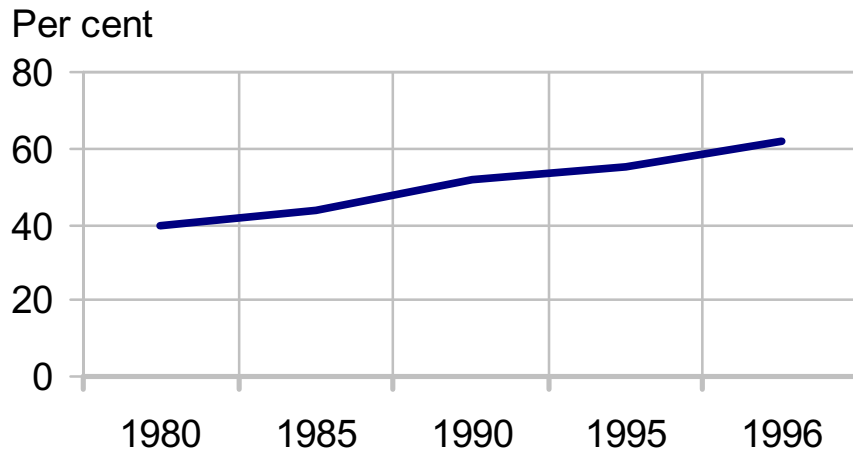


- When showing projections, the projection portion should be clearly differentiated graphically from the rest of the curve

Line charts/different types:

- Most often: periods or points in time (years)
- Area charts (accumulated line charts)
- Periodical data (months, quarters)
- Index
- Cumulative charts

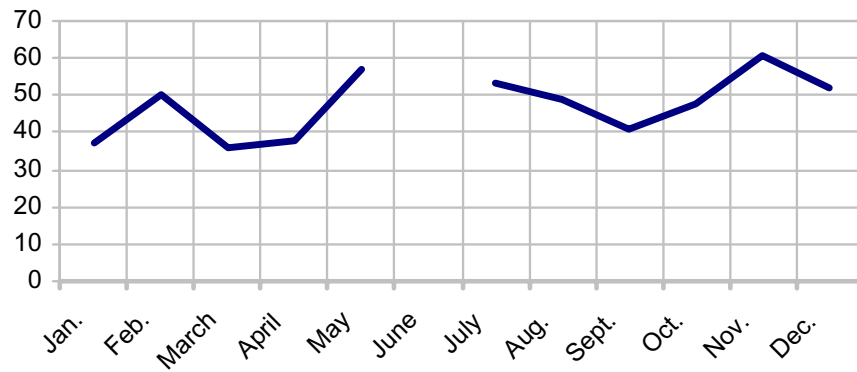
Standard line charts



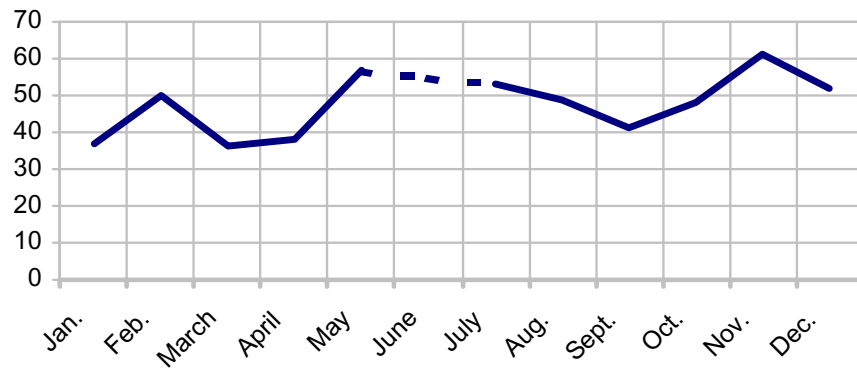
- “A time series” requires at least 4? data points (if not: use bar chart)

- Be careful when the data points are not evenly spaced; like 1980, 1985, 1990, 1995 and 1996

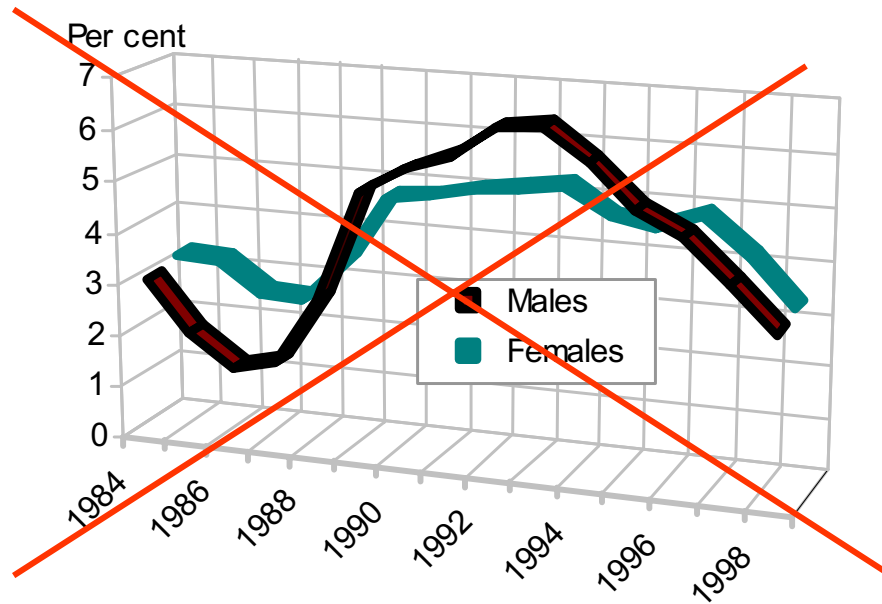
Missing data in time series



- The gap should be bridged with a dashed line

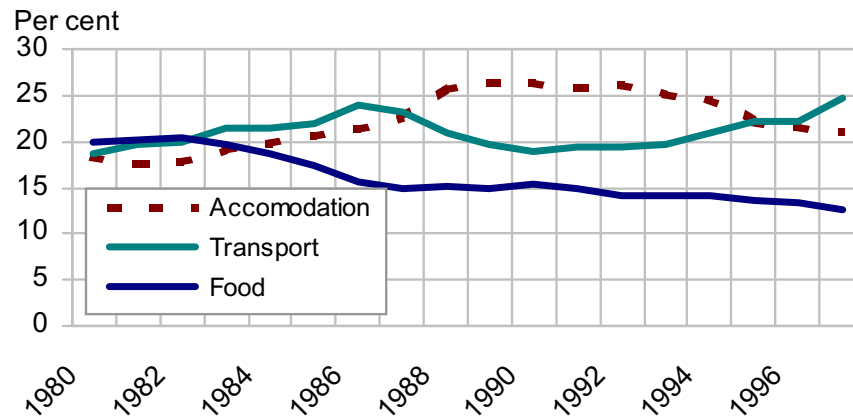


“3D” line charts?

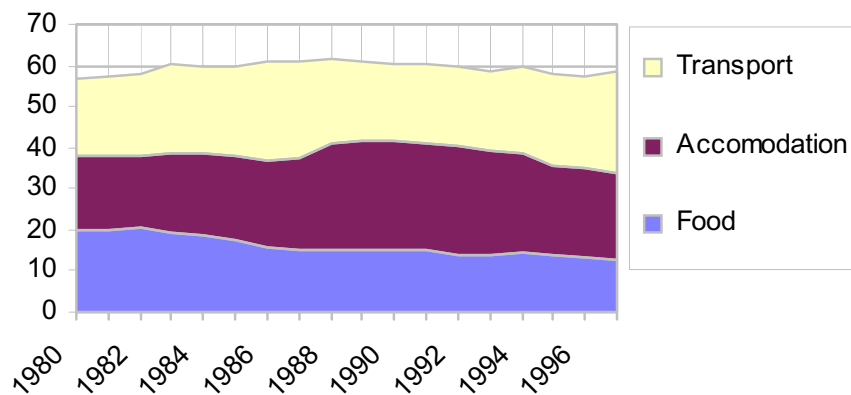


- **NO!!**

Area chart

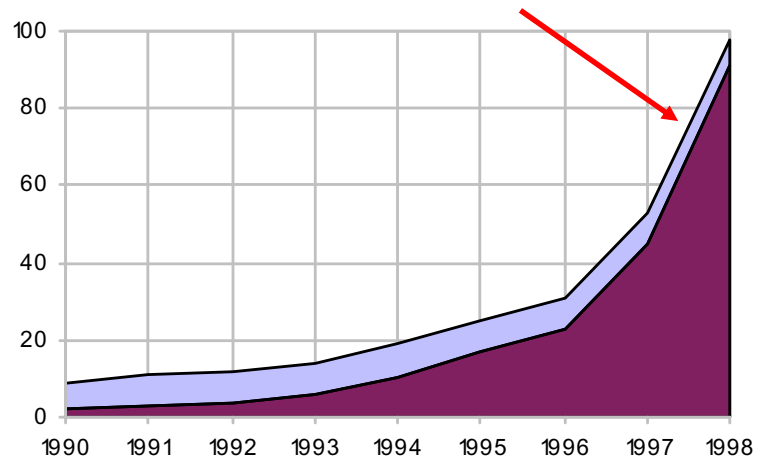


- Area charts are accumulated line charts; like stacked bars.

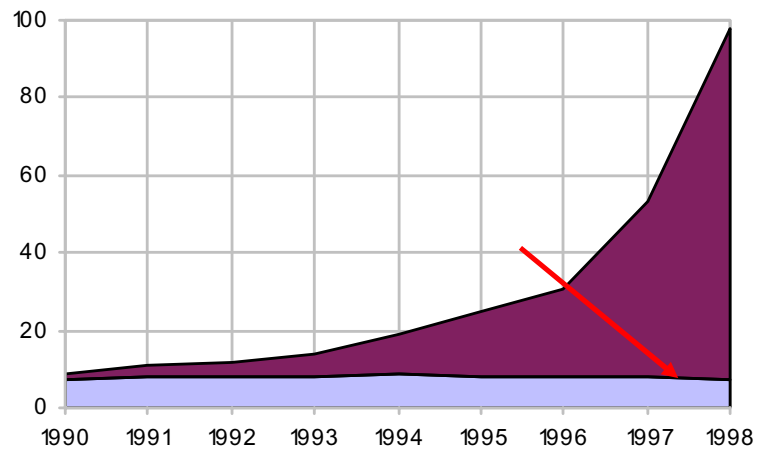


- Shows total and parts
- Don't use with too many groups/values

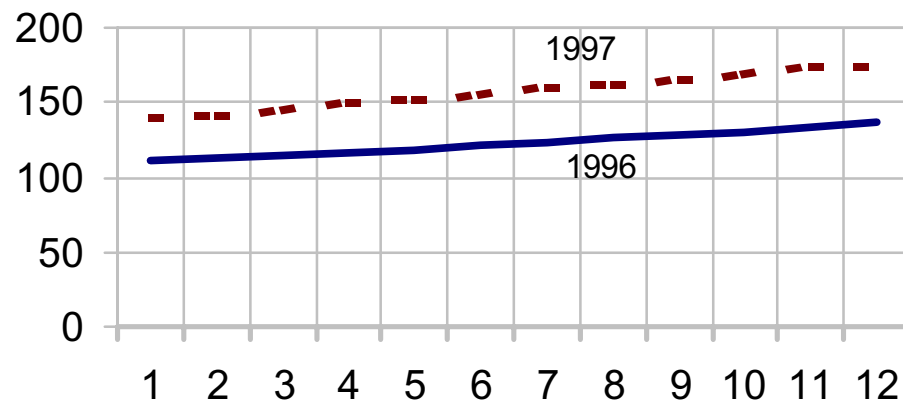
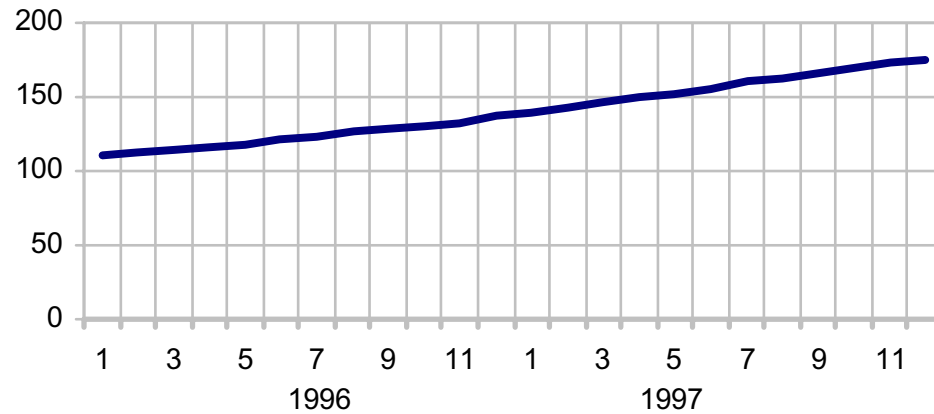
Area chart



- Order of series may be important

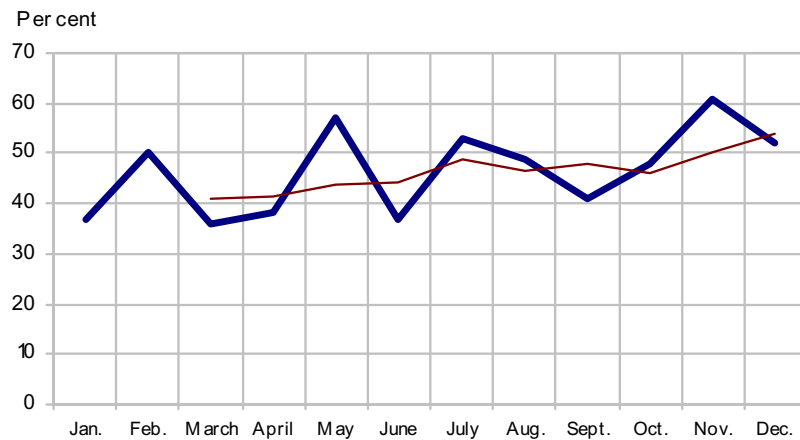


Periodical data



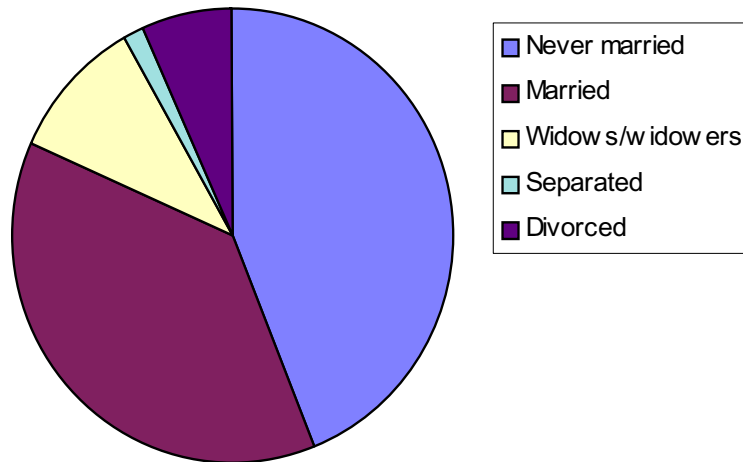
- Weekly, monthly and quarterly data. Monthly data for one year can be presented as a standard line chart
- With two or more years, we have two alternatives

Curve smoothing



- Often used with periodical data
- Moving average (here: 3 years)

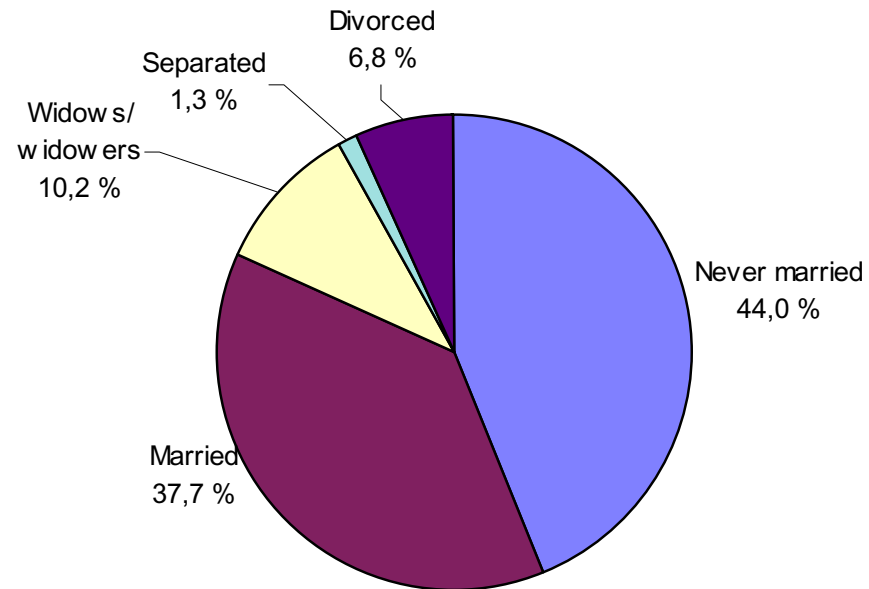
Pie chart



- Shows distribution of (qualitative) variables
- No axis
- Total area = 100%

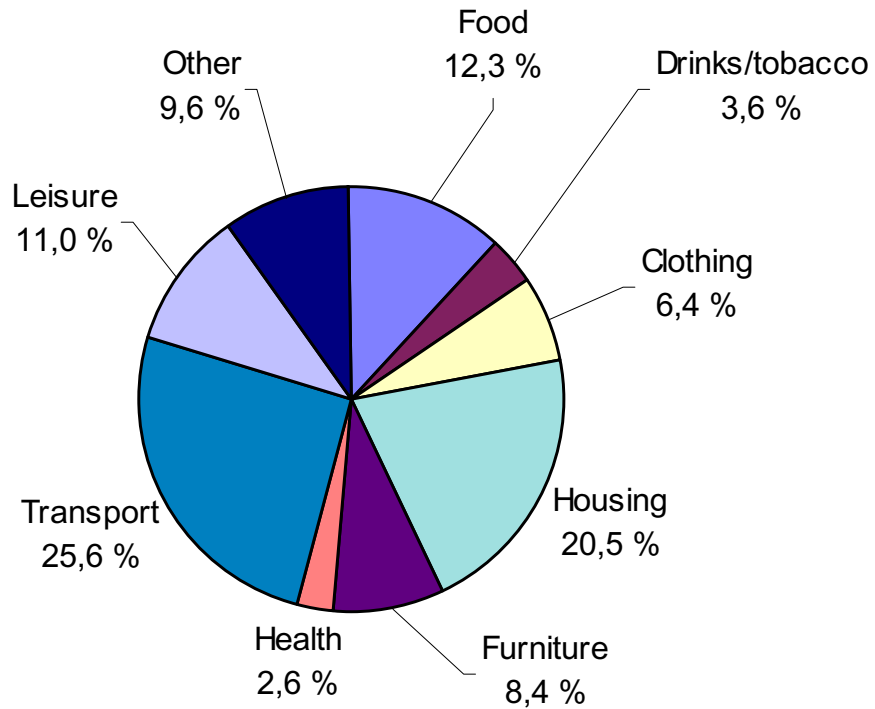
- Instead of legend....

Pie chart



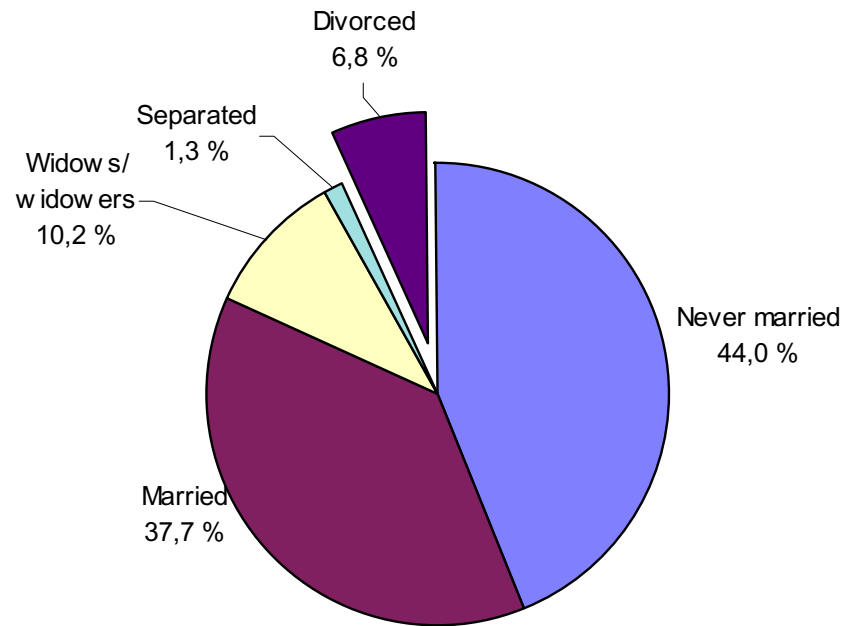
- ...use labels
- Since pie charts have no value axis, show percentage
- Maximum 5 values?

Pie chart



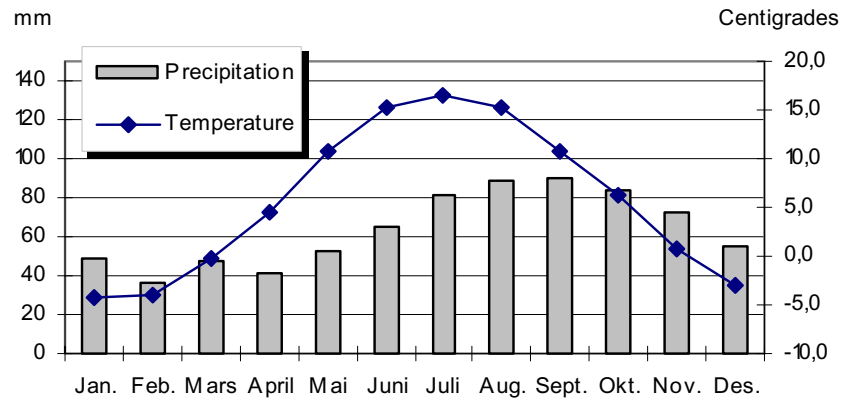
- With too many values, the reader loses interest

Pie chart



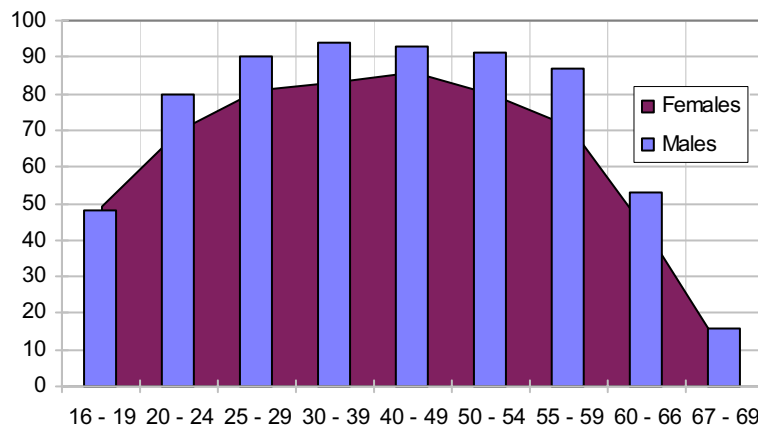
- If we want focus on or draw attention to a special group/sector, this can be done by “exploding”

Other types



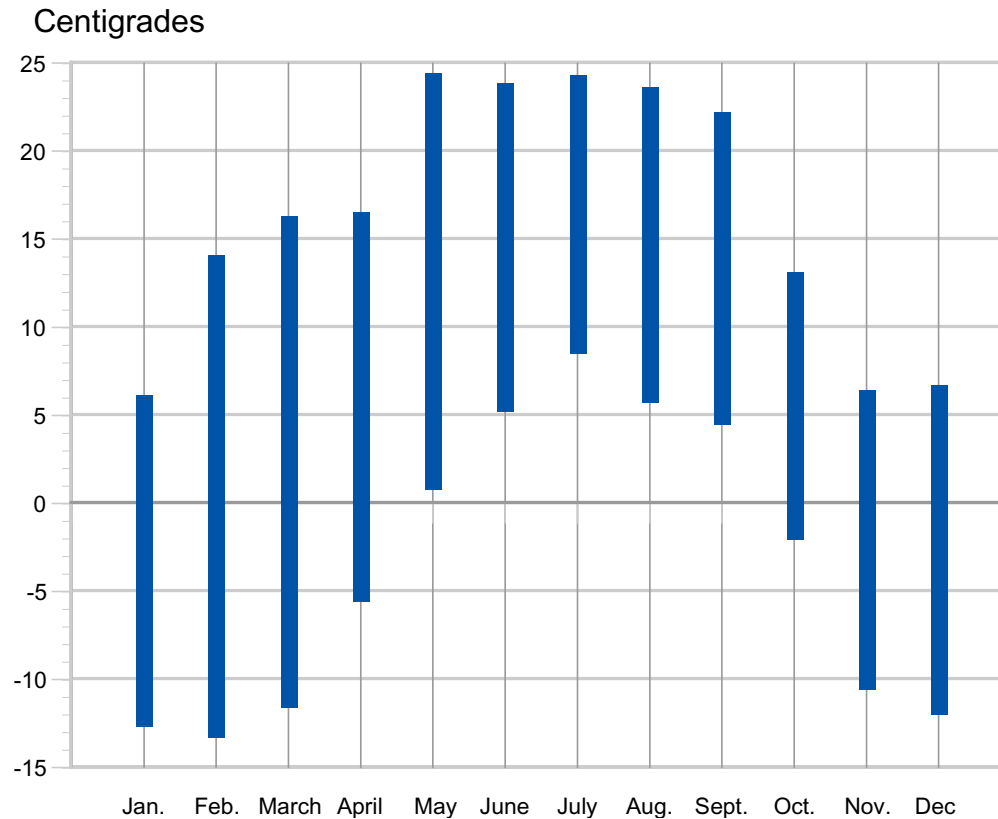
- Combination graphs (also called “overlay”):
 - Bar & line

Male and female labour participation rates by age. 1998



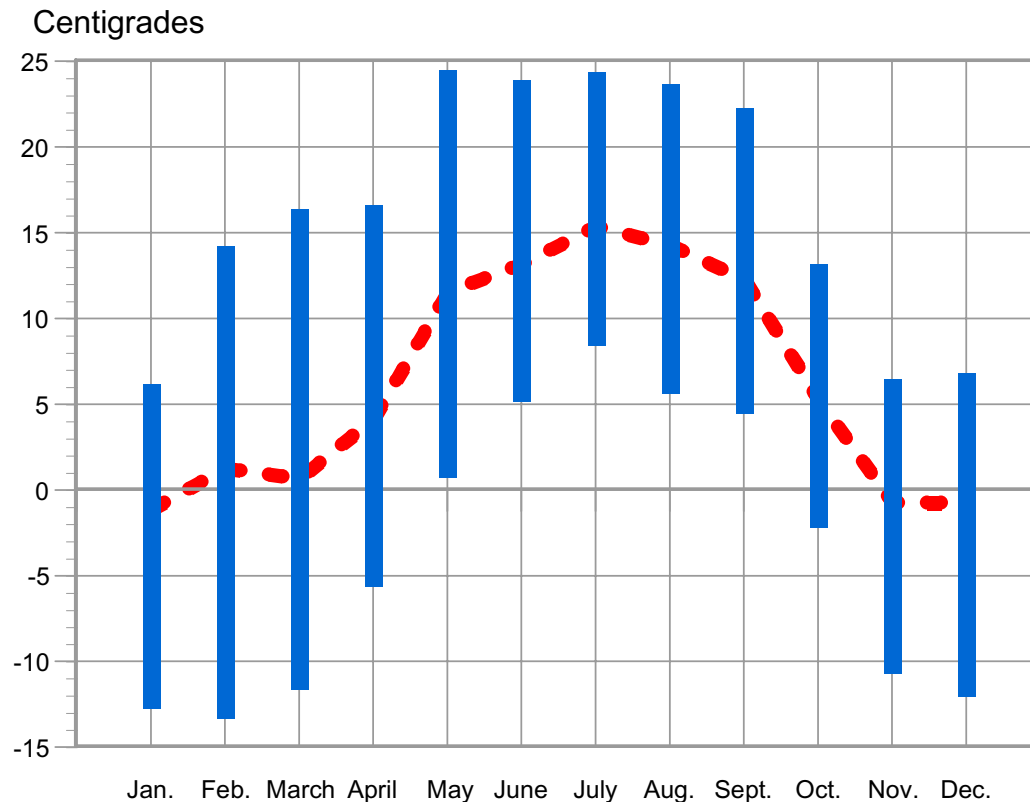
- Bar & area

Other types



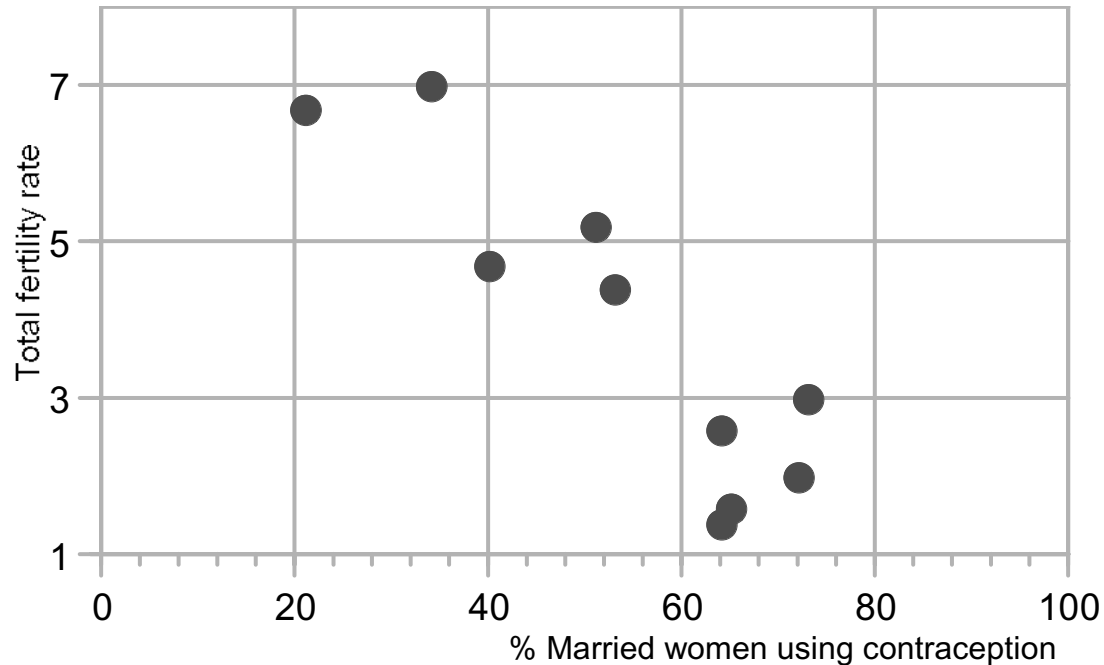
- Min. - max.
(floating column)

Other types



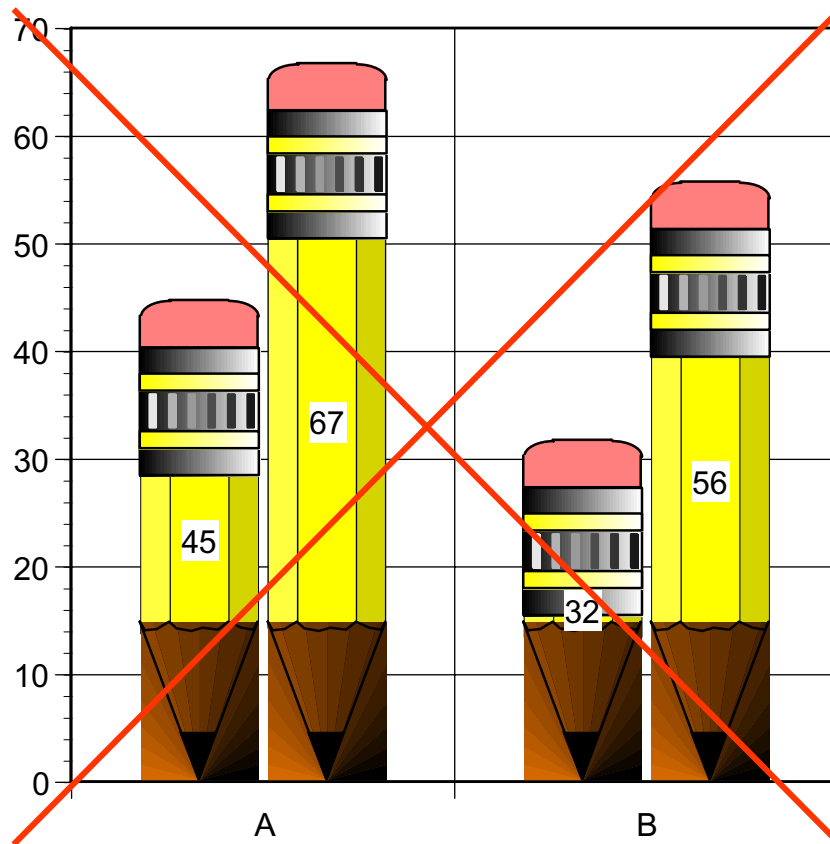
- Or: Min. - max.
+ average
temperature

Other types



- Correlation chart
- have two value axes
 - show the relationship between variables.
 - Often used with regional data (municipalities, regions, countries)

Other types



- Pictogram: Not recommended