Get every one in the picture

Graphing your data

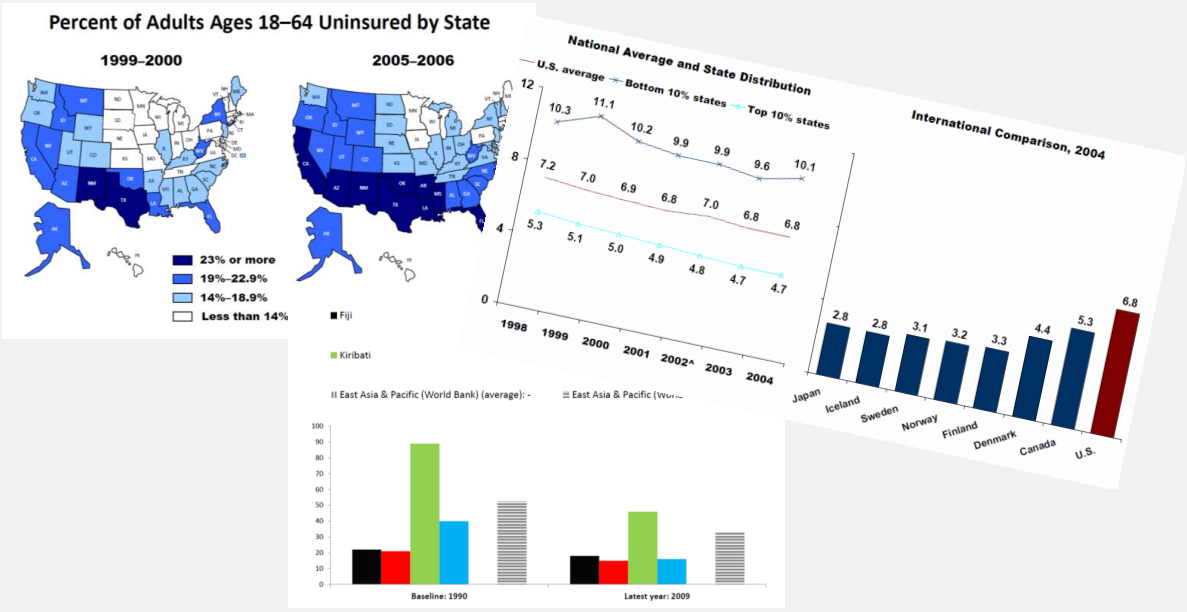
Data analysis and Report writing workshop for Civil registration and vital statistics data.

UNITED NATIONS ESCAP
Economic and Social Commission for Asia and the Pacific

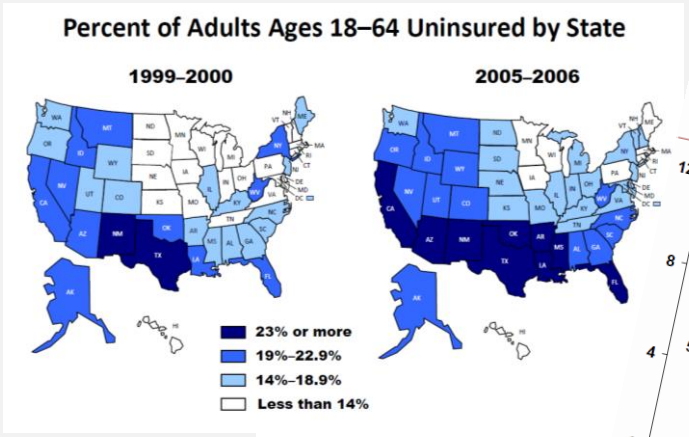
Adapted from Pacific Community's
Data analysis and report writing
Workshop for the North Pacific

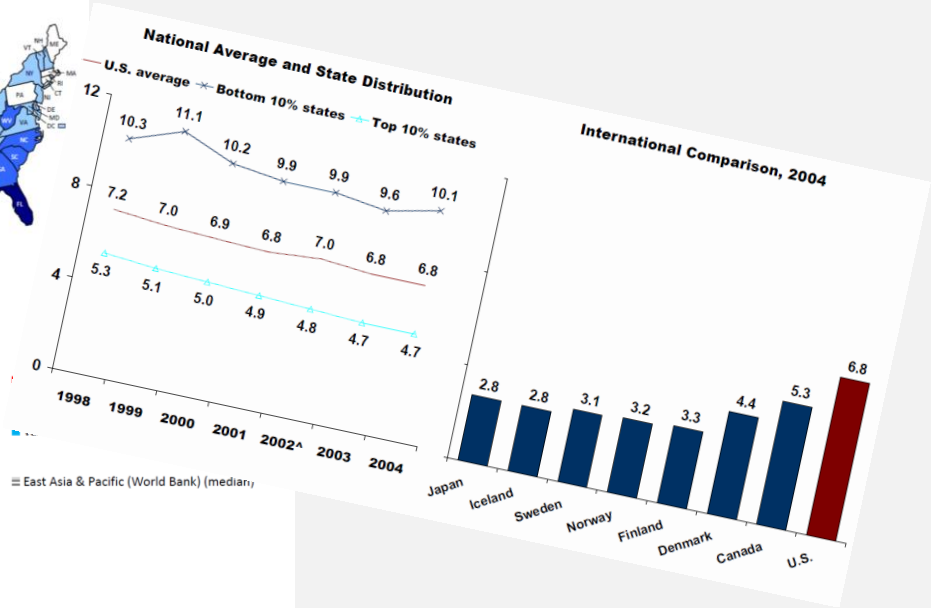
Why use graphs?

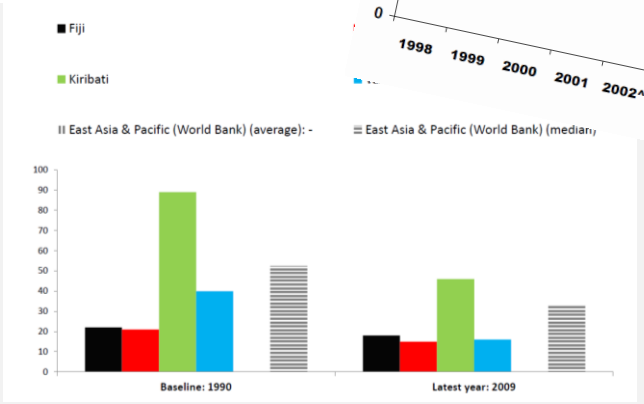
A picture is indeed worth a thousand words, or a thousand data points, and is extremely effective in expressing key results.



Why use graphs?







Comparison
Change over time
Distribution of items
Relation

Three key guidelines for any graph

1. **Who is your target audience?** What do they want to know about the issue?

2. **What message do you want to communicate?** What do the data show? Is there more than one message?

3. **What is your message?** Do you want to compare items, show time trends, or analyse relationships in the data?

Different graphs show different aspects of the same data

A good graph

- Grab’s attention
- Presents information clearly, simply, and accurately
- Does not mislead
- Displays data in a concentrated way
- Facilitates comparisons
- Highlights trends
- Illustrates messages

General rules for preparing graphs

- Let the data determine the type of the graph.
- Be simple, not load too many things into a graph.
- Graph should have a clear, self-explanatory title.
- Include the source of the data.
- The units of measurement should be stated.
- All axes should be carefully labelled.
- The scale on each axis should not distort or hide any information, i.e. the graph should show the data without changing the message of the data.

General rules for preparing graphs

- Graphs should clearly show any trend or differences.
- Start Y axis scale at zero or use scale breaks to avoid misinterpretation (exception life-expectancy graph starts at 40)
- Graphs should be accurate in a visual sense
- Make all text on graph easy to understand:
  - Avoid abbreviations/acronyms
  - Write labels from left to right
  - Use proper grammar
  - Avoid legend except on maps

General rules for preparing graphs

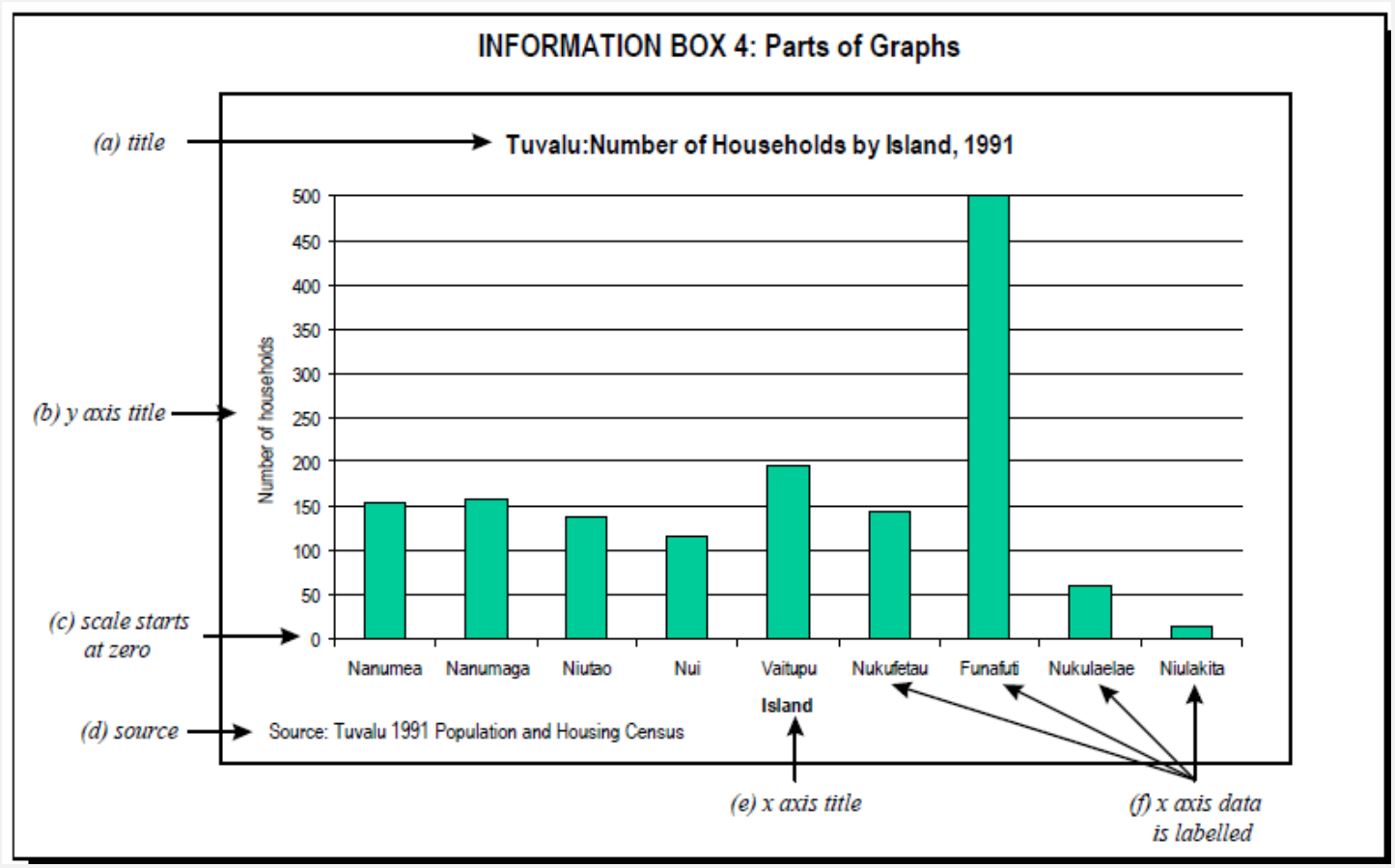
- Choose type of presentation carefully:
  - Size and shape of plotting symbol
  - Method of connecting points
  - Use two-dimensional design for two-dimensional data
  - Use solids rather than patterns for line styles and fills
  - Avoid data point markers on line graphs

- When comparing different populations, keep the same scale.

Components of a good graph

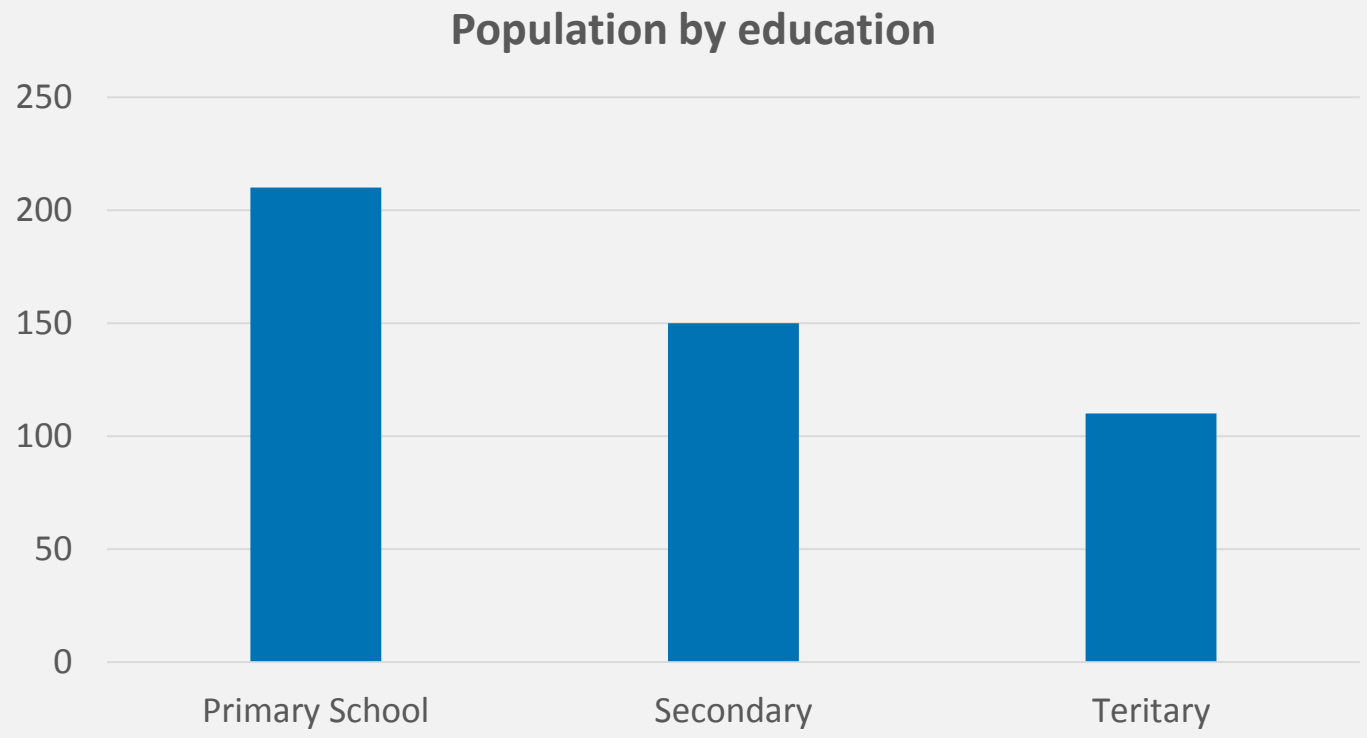
1. **Data components**
   - Bars, lines, areas or points
2. **Support components**
   - Chart title
   - Axis labels
   - Axis titles
   - Gridlines
   - Legend and data labels
   - Footnote
   - Data source
3. **Decorative features**

Components of a good graph



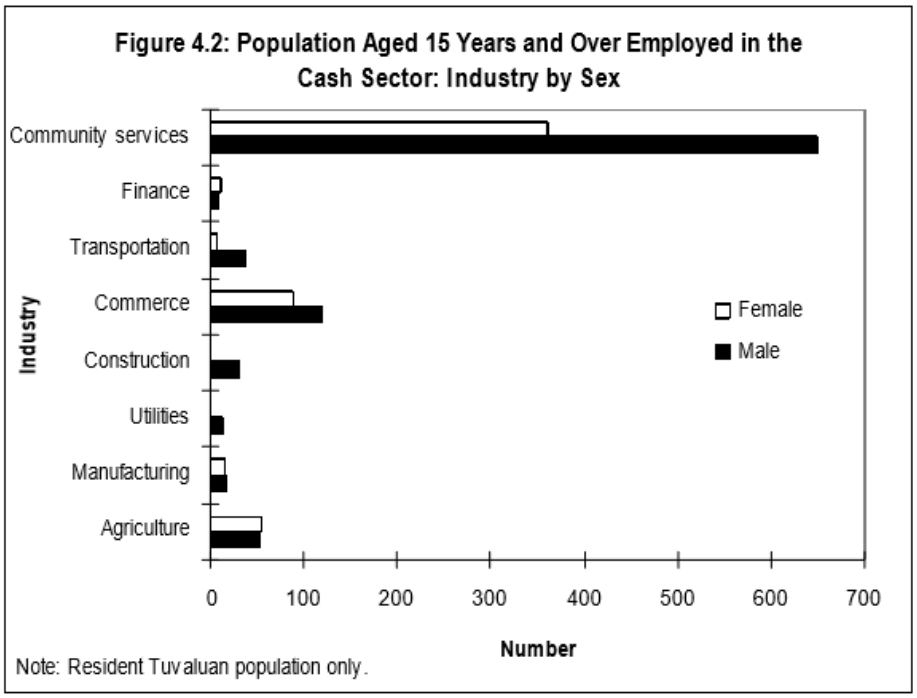
Which type of graph?

Bar charts



- **Used to describe the distribution of qualitative and quantitative data (grouped into equal sized class intervals) and to compare groups**
- **Often ordered from largest to smallest**
- **Can be vertical and horizontal.**

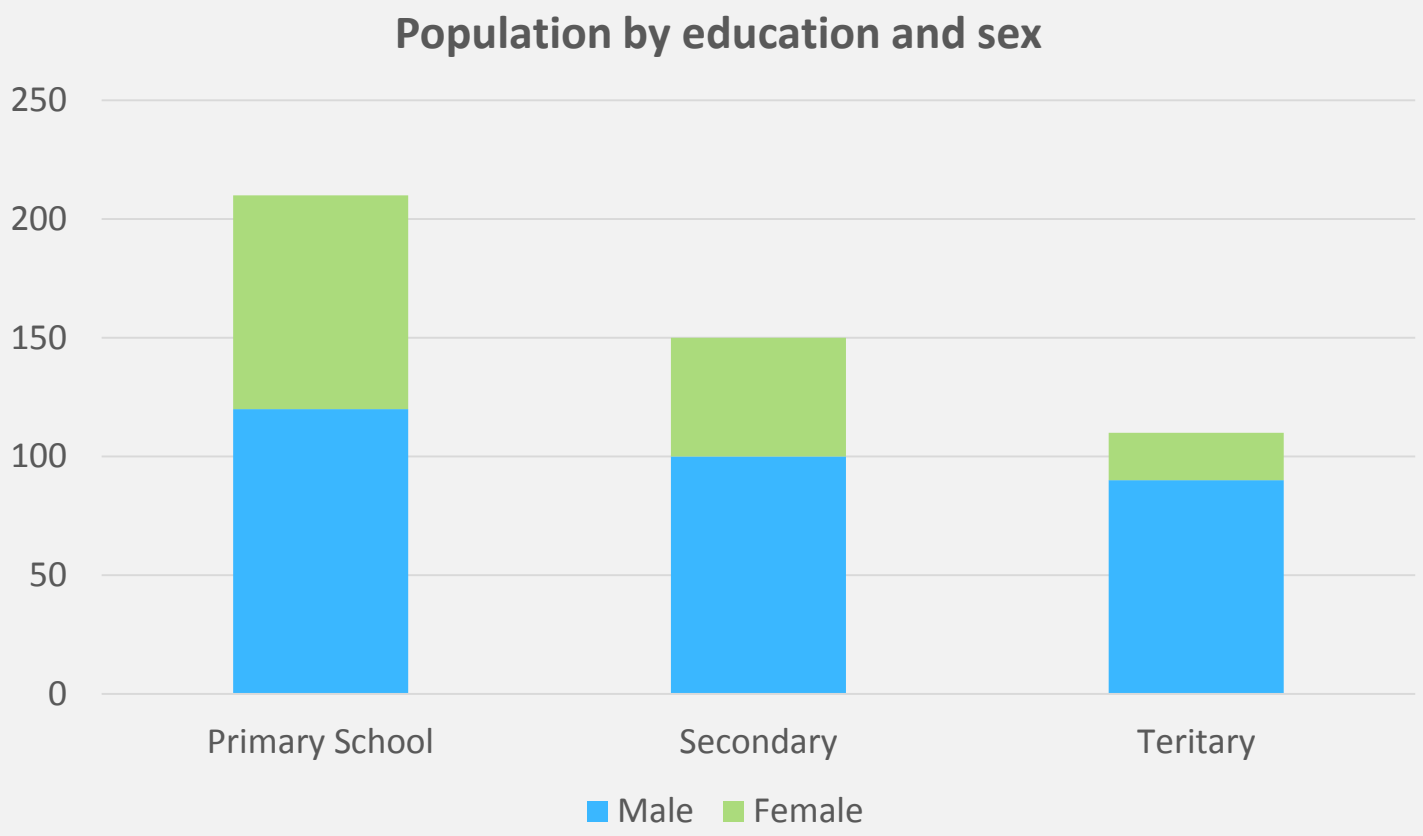
Bar charts



Source: Tuvalu 1991 Population and Housing Census.

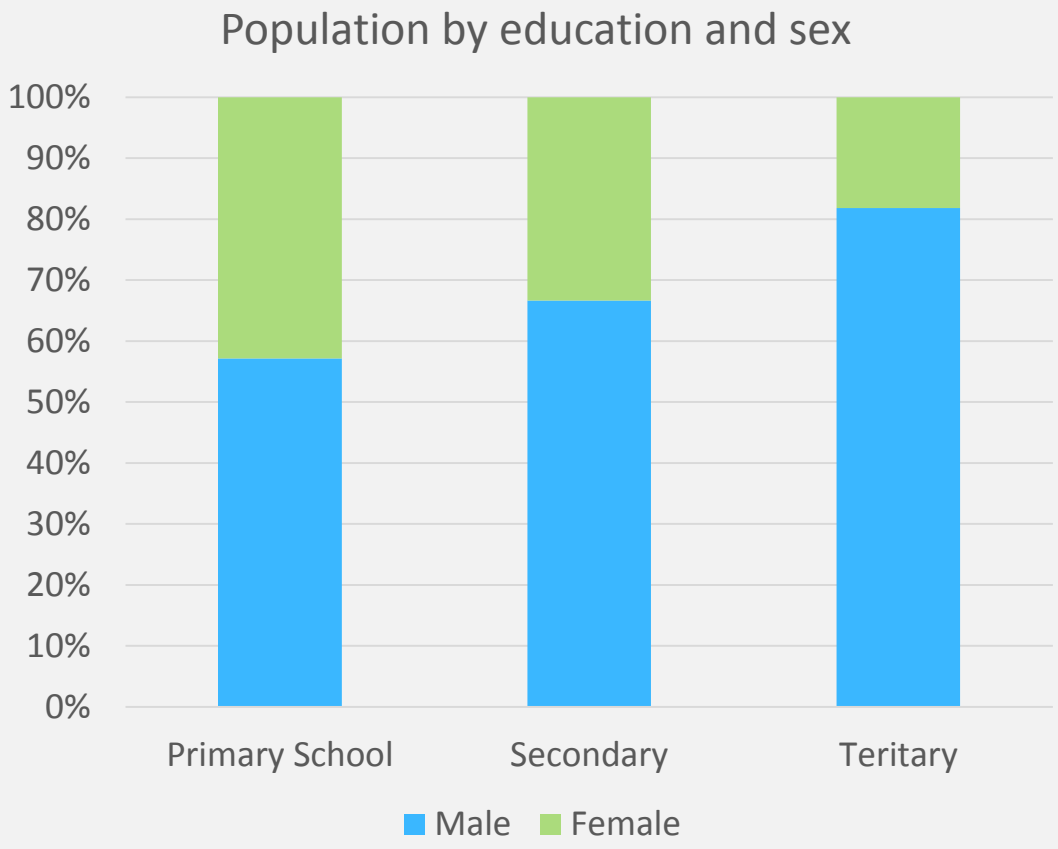
- **Horizontal bar charts are used when the labels for the variable categories are too long to fit neatly in vertical chart**

Stacked bar charts



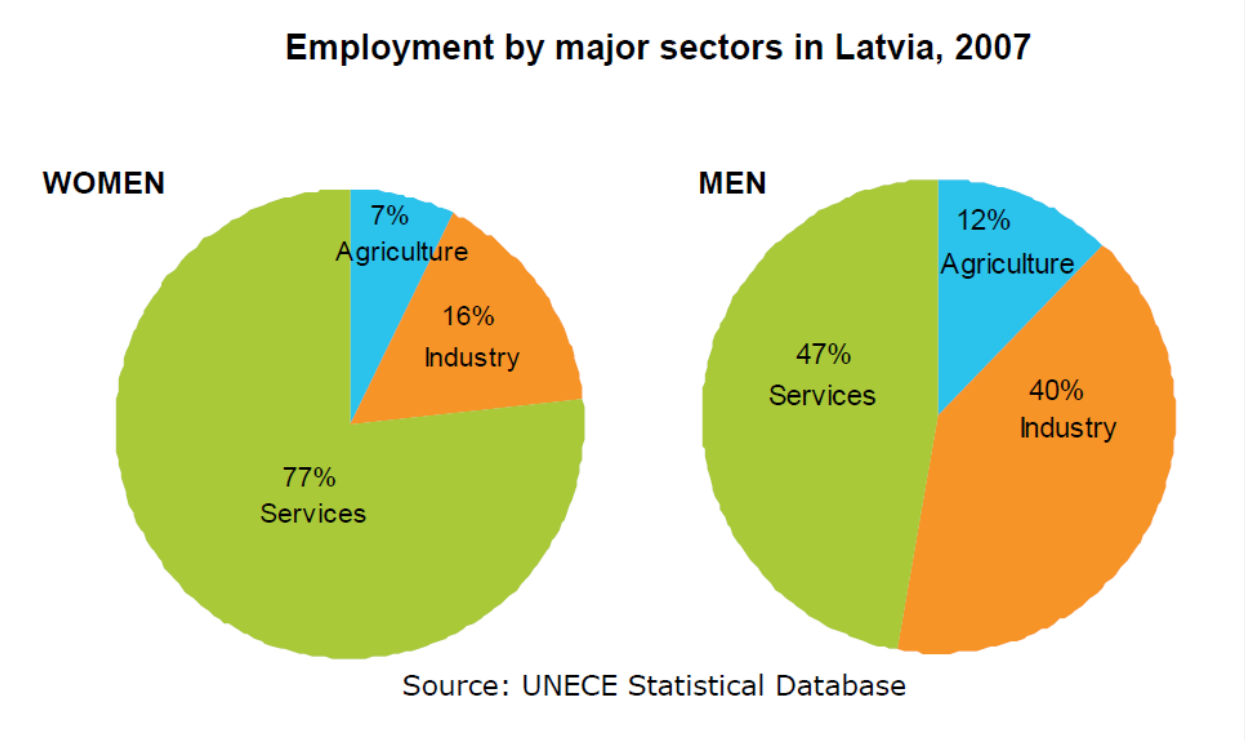
- **Shows and compares segments of totals**
- **Not suitable if too many items in each stack**
- **Difficult when items are fairy close in size**

100% Stacked bar charts



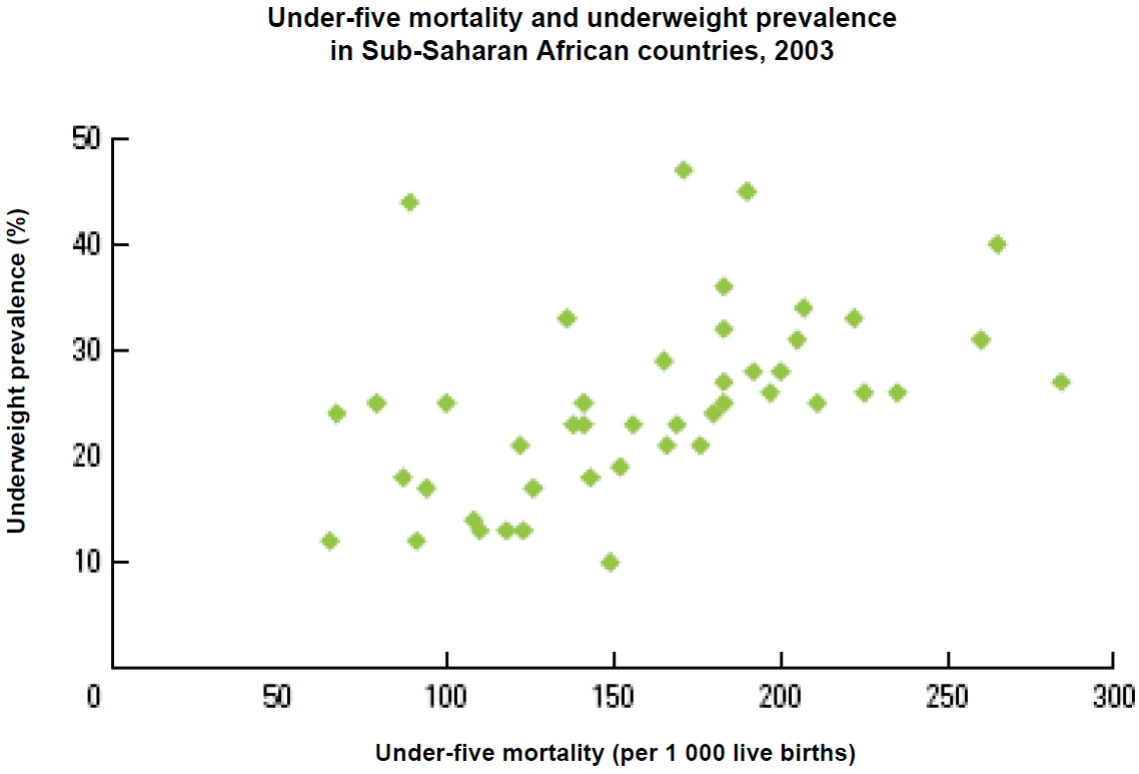
- **Shows and compares percent of segments of totals**

Pie charts



- **Used for qualitative data**
- **Each slice represents the relative frequency of each category**
- **Most effective with five or fewer classes in the data**
- **Should be ordered by size**
- **Not always easy to compare within and between categories**

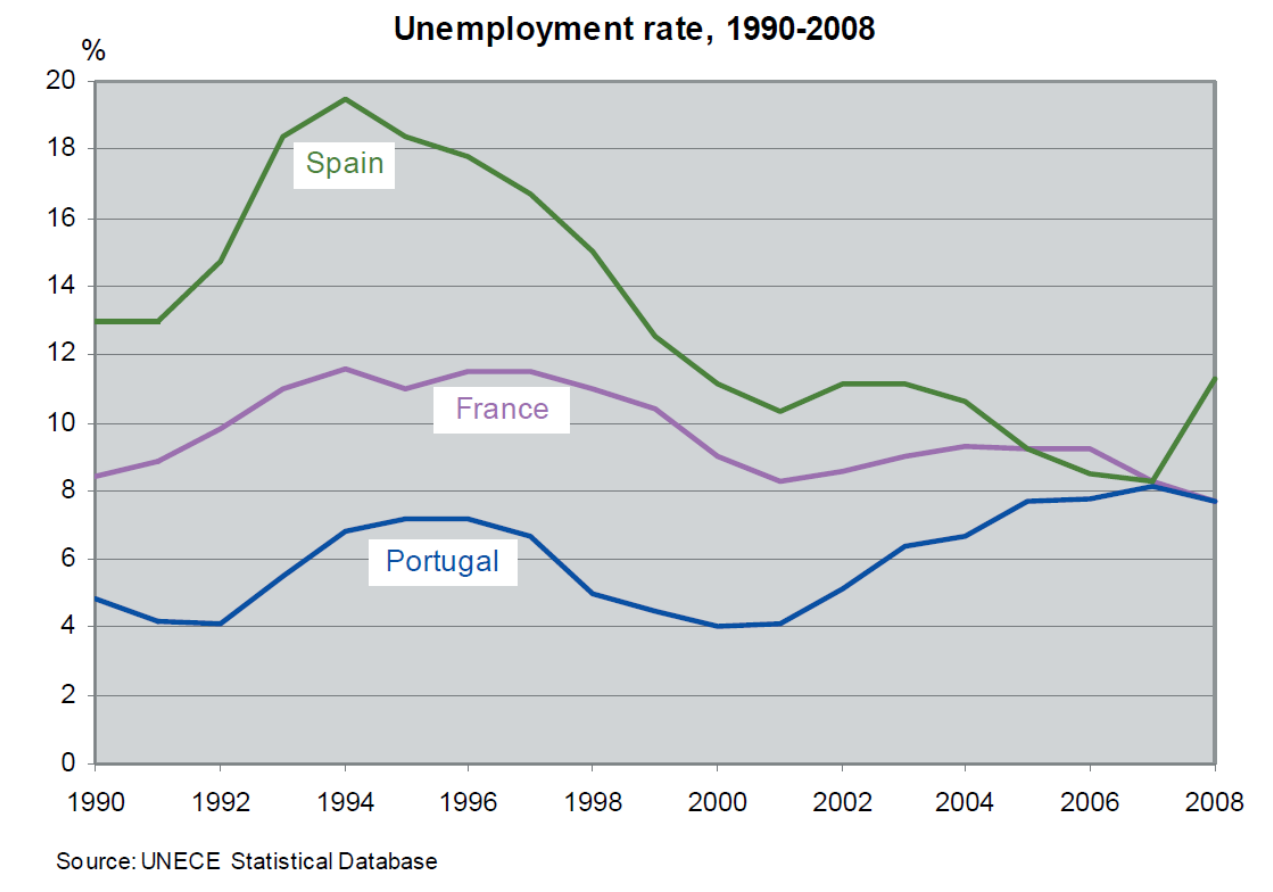
Scatter plots



Source: Jamison et al. (2006) *Disease and Mortality in Sub-Saharan Africa, 2nd edition*, Washington D.C., The World Bank[7].

- **Simple to see relationships clearly**
- **Can be used to detect outliers**

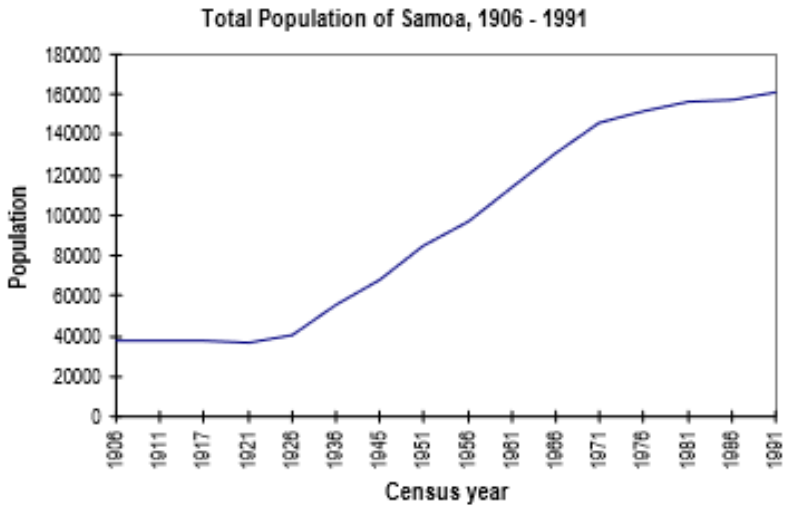
Line graphs

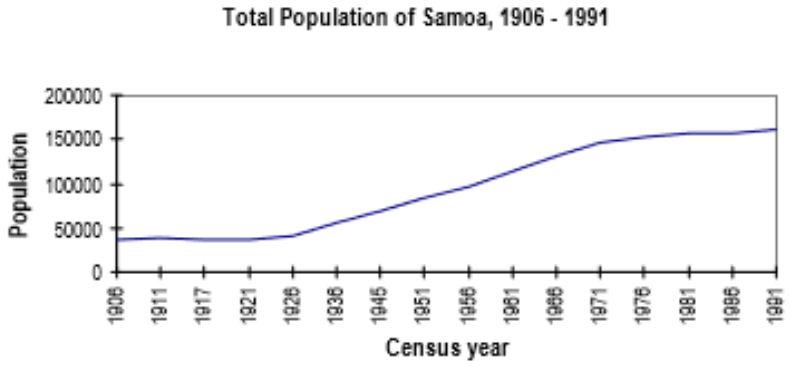


- **Used for bivariate data, (independent variable on x-axis, dependent variable on y-axis)**
- **Commonly used to display trends over time**

Line graphs

Figure 4.4: Total Population of Samoa, 1906 – 1991

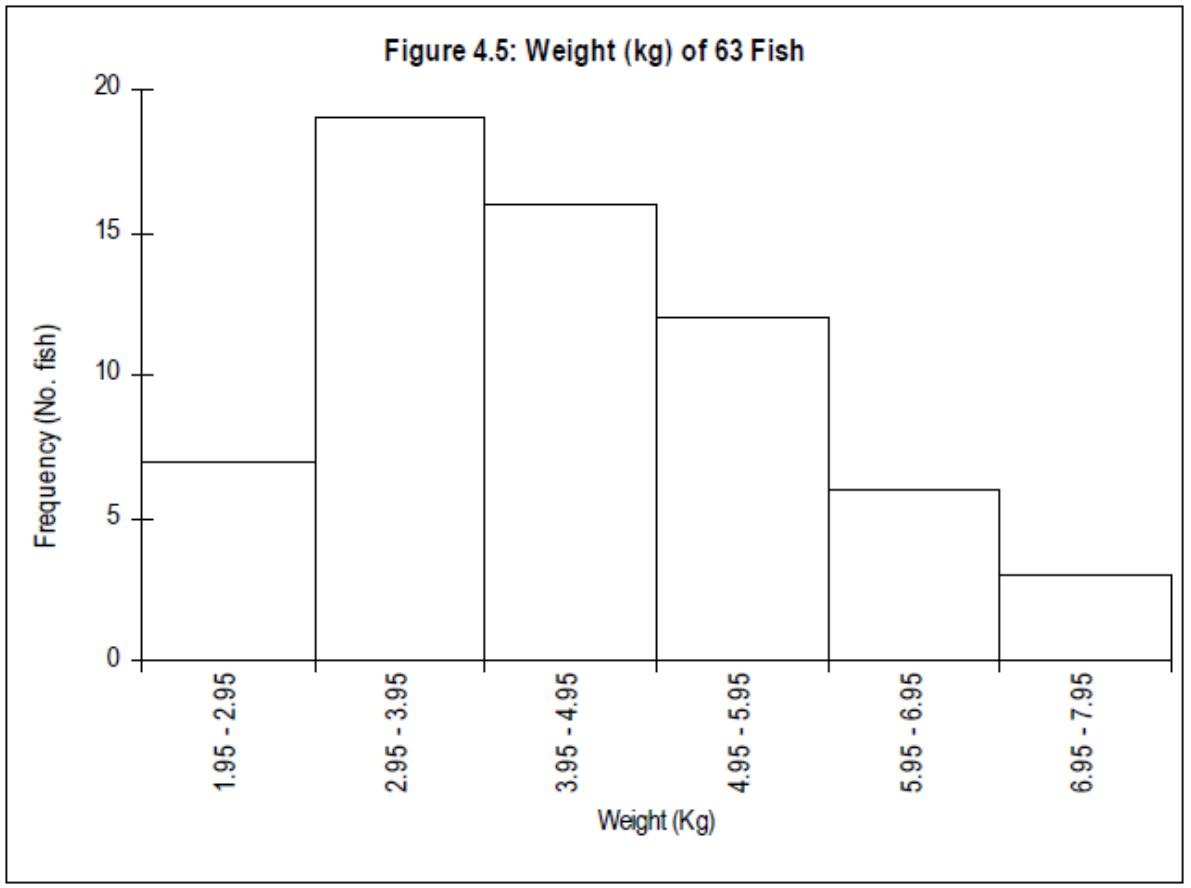




Source: Census of Population and Housing, 1991, Western Samoa

- **Be careful! Scale on the y-axis should not give wrong impression**

Histograms



- **Common method of representing a frequency distribution for continuous data.**
- **Area of the bars is proportional to the class frequencies.**

Difference between histogram and bar charts

- In histogram area of the bars (not necessarily the height) is proportional to the class frequencies.

- In histogram, bars are always side by side, without gap, reflecting the continuous nature of the data.

- Generally histogram has equal width bars.

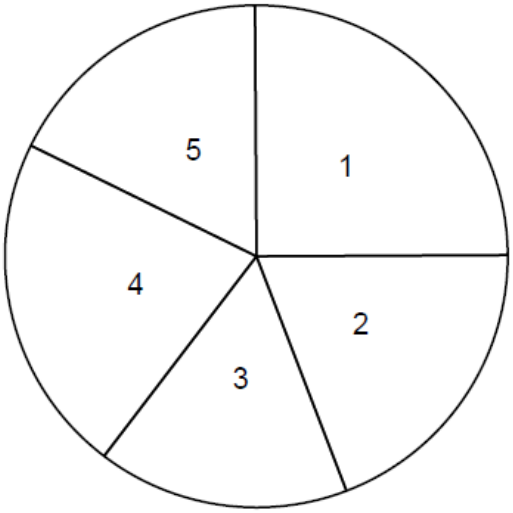
Comparing different charts

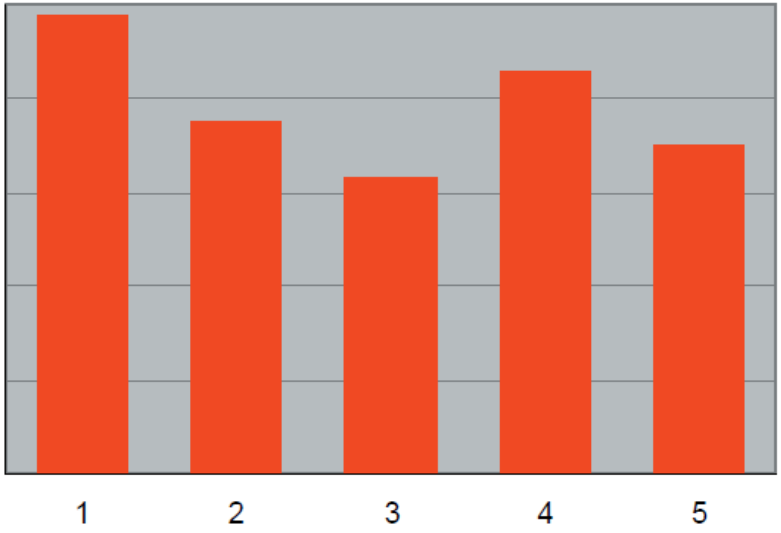
Chart type	Characteristics
Bar chart (vertical)	Simple and clear Works for categories and time series Not good for long time series Small space for long names
Bar chart (horizontal)	Good for large number of categories Works for long names Not appropriate for time series
Pie chart	Used for qualitative data Not good for making comparisons Too many ‘slices’ gets confusing
Scatter plots	Shows relationships between variables Used to detect outliers Can be difficult to interpret
Line chart	Simple and clear Best for time series and trends More than three lines gets confusing

Which type of graph?

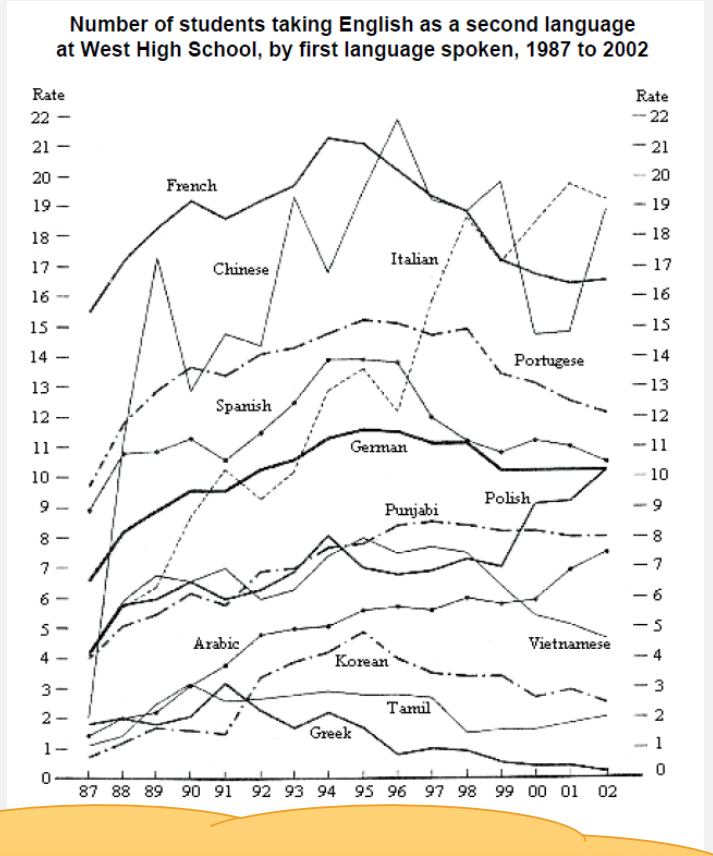
- A trial and error approach can be very helpful.
- Check which type communicates the message best.

Which type should I use?



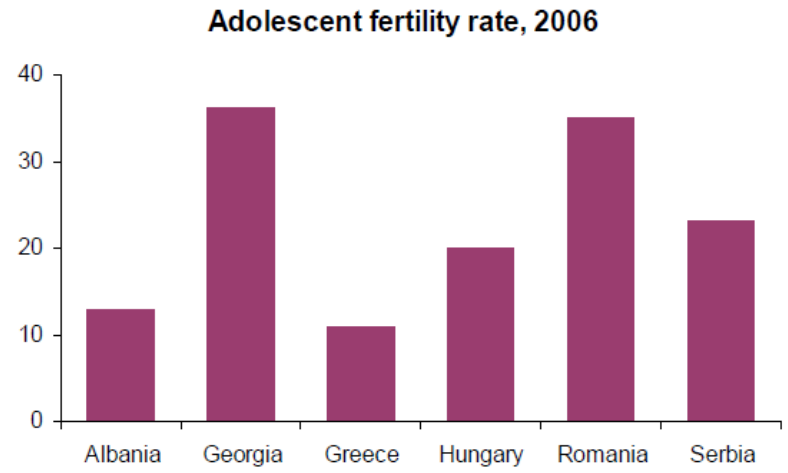
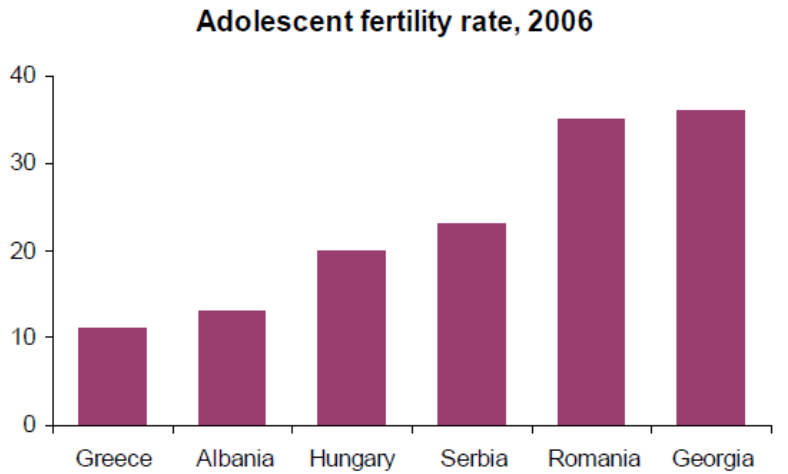


No more bad graphs!

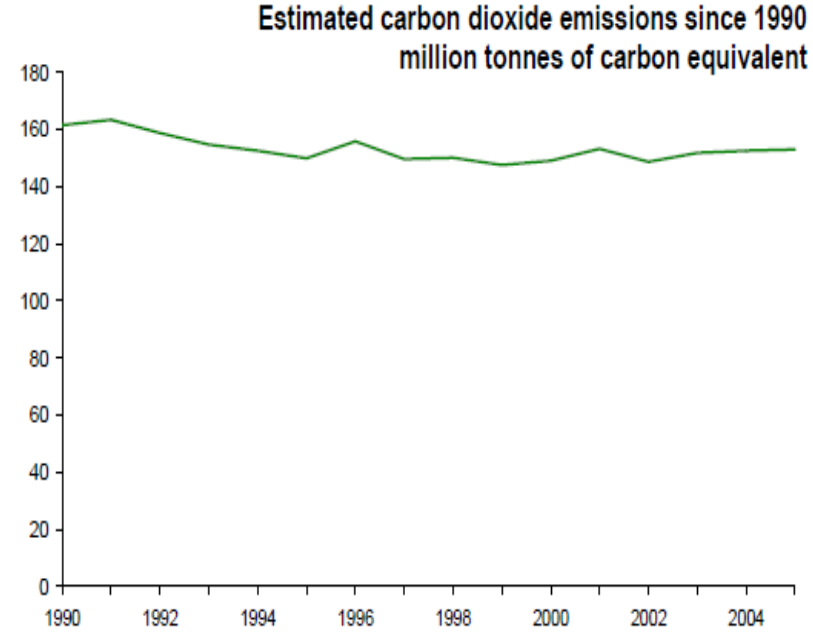


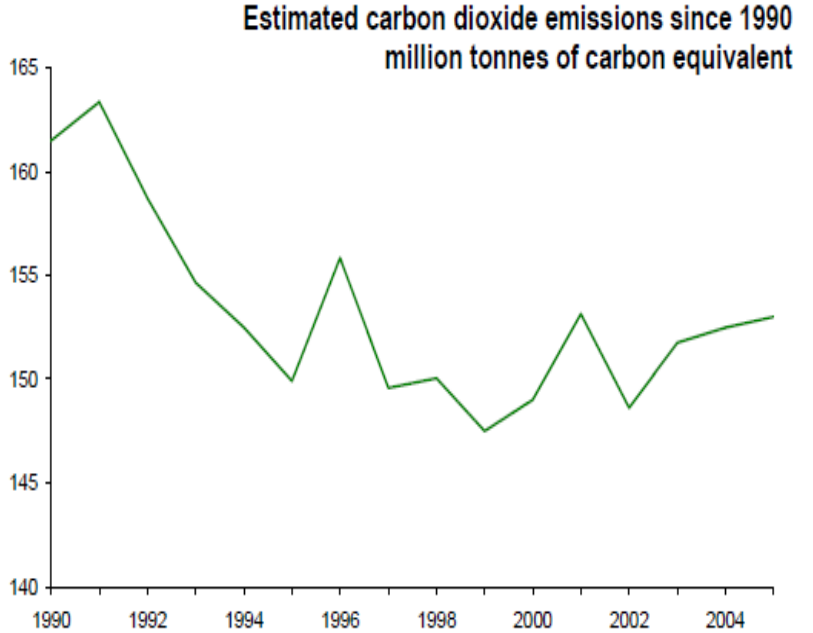
Complication fails to communicate

Sort your data

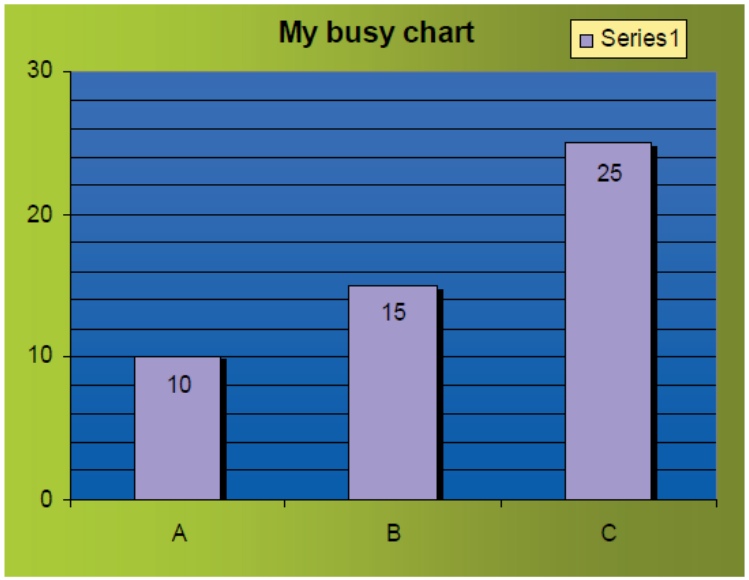
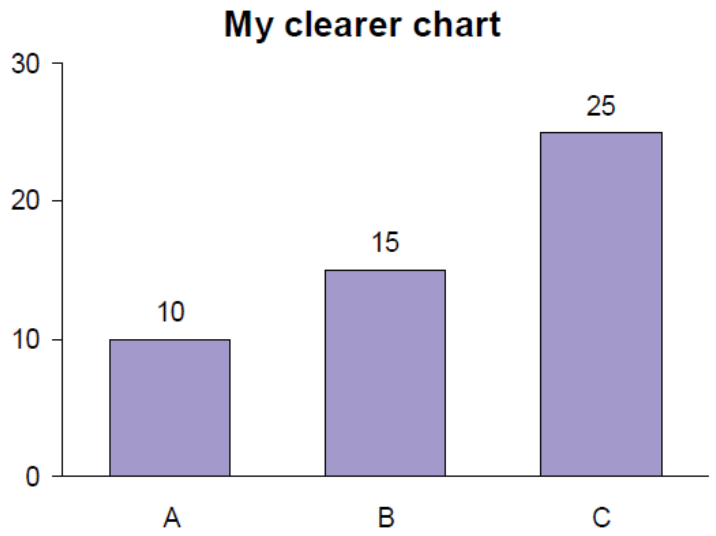
BAD EXAMPLE	GOOD EXAMPLE
	
The data are presented by alphabetical order of countries. The values are very difficult to compare. Attention is on the first and last values, which have no specific relevance.	The data are presented in order from smallest to largest values. It is easy to compare them. Attention is focused on the minimum and maximum values of the dataset.

Remember to be accurate



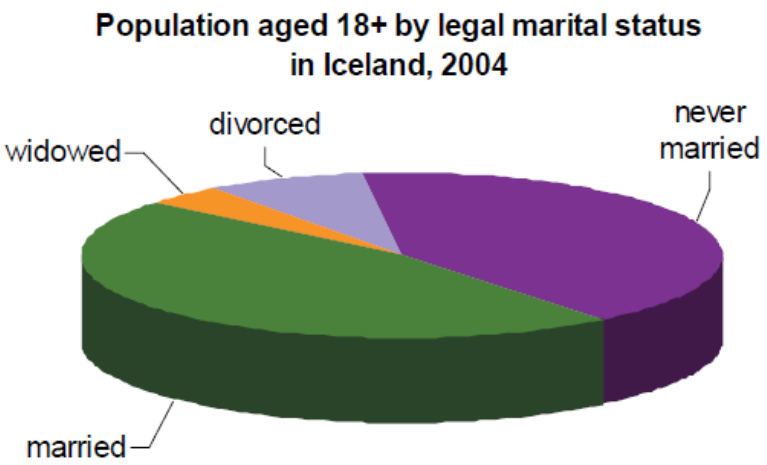


Keep it simple

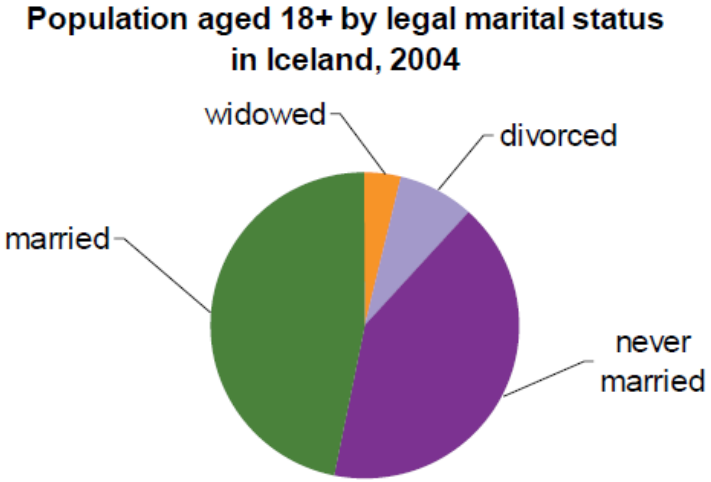
BAD EXAMPLE	GOOD EXAMPLE
	
All components have maximum impact. The result is a busy chart, difficult to read, even though it shows only three values.	This chart is much easier to read. Minimal use of support components ensures that data take centre stage.

Keep it simple

BAD EXAMPLE

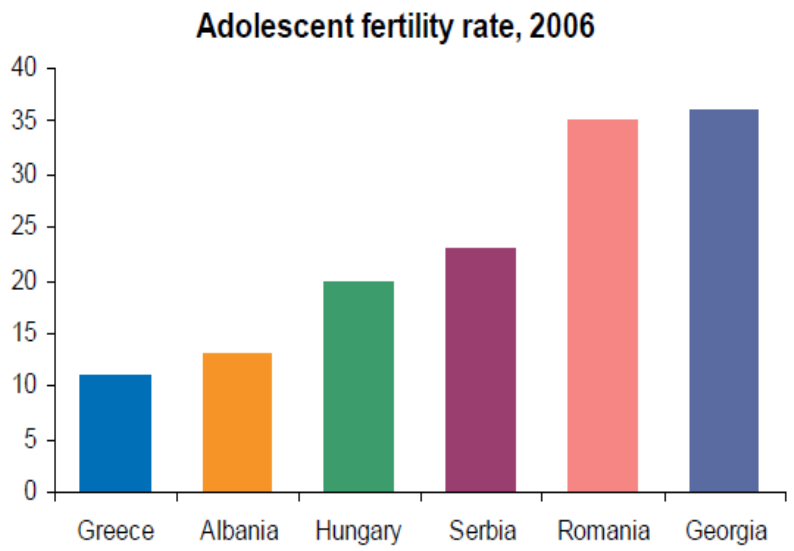


GOOD EXAMPLE

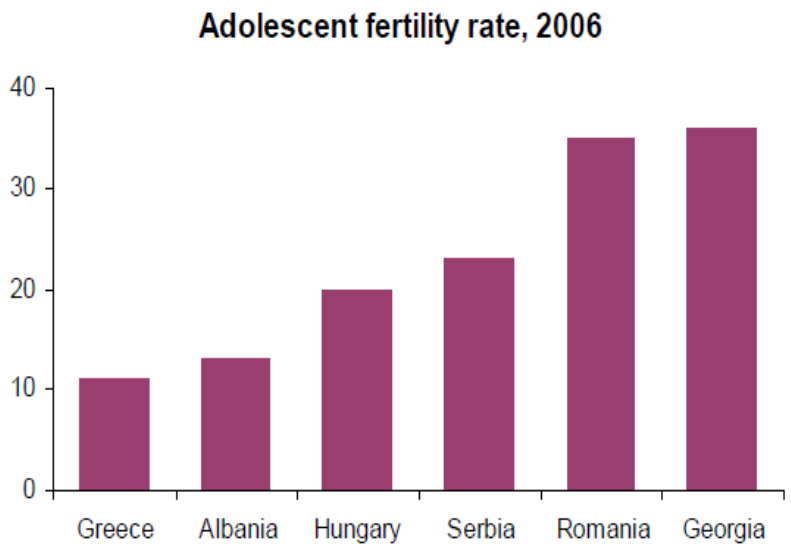


Keep it simple

BAD EXAMPLE



GOOD EXAMPLE



Graphing in Excel

- The main thing is to select the correct x and y variables.
- Select the cells you want to chart
- Under INSERT tab, select the type of chart you want to create
- To edit the chart, click on your graph and chart tools options
- Use the Design tab, and Format tab change chart type and layout
- To change the look of something on your chart, select the object and double click on it.

Graphing in Excel

- Make some charts for your test data and country data for percent distribution of:
  - Birth by sex/weight-group
  - births by age/education of mother
  - births by place of birth
  - Attendant at birth (for your country data)

- Make a chart to check “does mother’s education have effect on baby’s weight”

- Graph your ASFRs for your test data and country data
  - In which age group does fertility peak?
  - Describe the pattern you see
  - How does adolescent fertility compare to the rest of the world?
  - For your country data, how do ASFRs change over time or between regions?