

Session 7.3

Presenting Data

Dr Annabel Bradbury, TRL Ltd

1 Data aim

2 When to use tables and graphs

3 Principles for tables

4 Principles for graphs

5 What graph for my data?

- 1 Data aim
- 2 When to use tables and graphs
- 3 Principles for tables
- 4 Principles for graphs
- 5 What graph for my data?

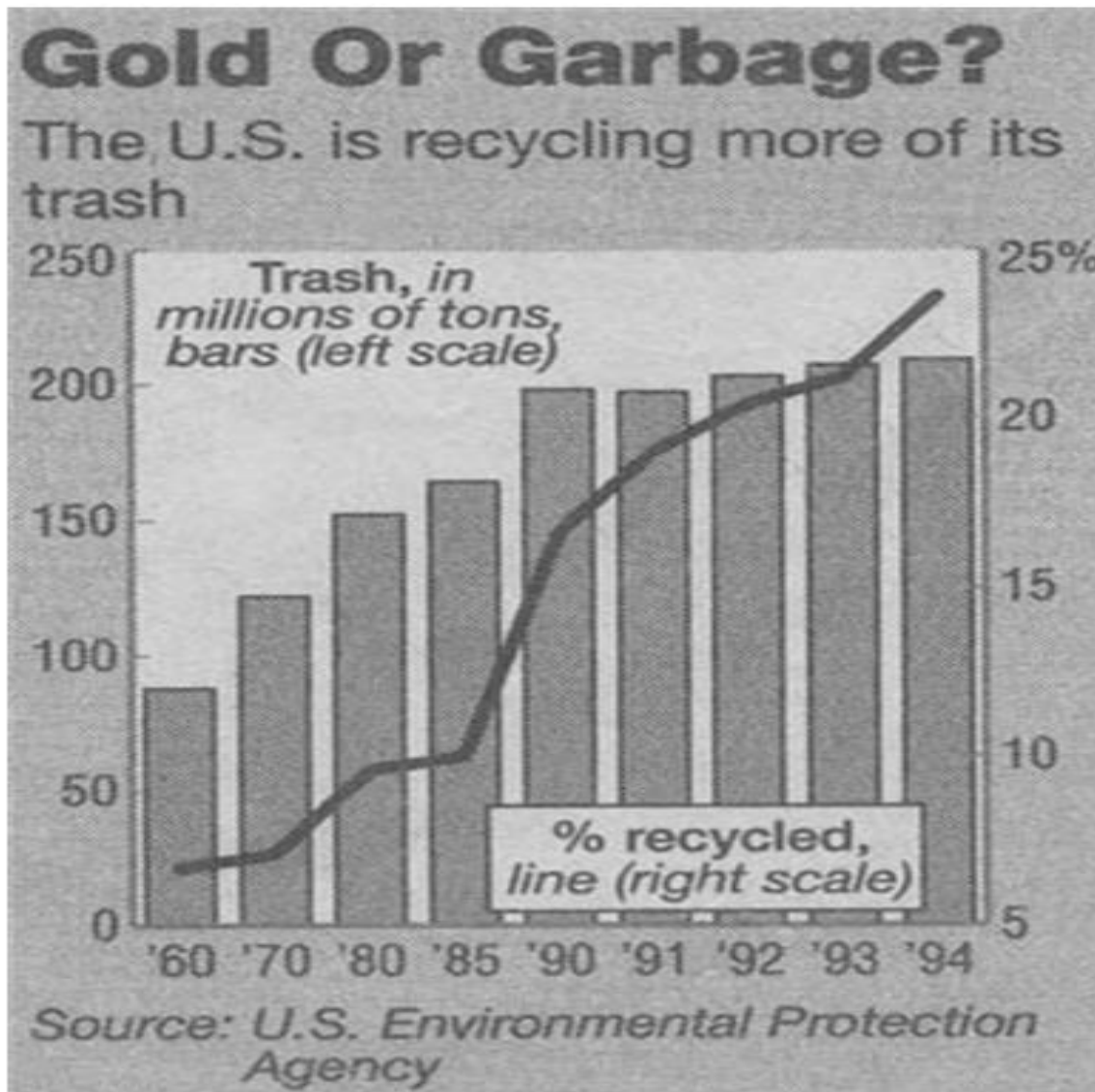
1: Writing an aim

- *Think about your audience*
- What do you want your audience to understand from the table/figure?
- Is it simple and useful?
- *How might your audience misinterpret your table/figure?*

Write an aim for your table or graph before you start building it.

e.g. I want my table to show relative amounts of mileage travelled by different vehicles in a year.

Audience: summary report to Ministry of Transport (published for the general public)



1 Data aim

2 When to use tables and graphs

3 Principles for tables

4 Principles for graphs

5 What graph for my data?

- Aim of statistical analysis, and data used, will influence choice of data presentation i.e. tables or a graph
- For displaying data use a table if:
 - The trend is not important (or at least not the most important feature); or
 - The number of values is small.

	2008	2007
LGVs (billion veh-km)	68.097	65
Motorcycles (million veh-km)	5141.5	5150
HGVs: 2 axles rigid	10.74	11
HGVs: 3 axles rigid	2.0122	1.6
HGVs: 4 or more axles rigid	1.8536	1.9
HGVs: 3 and 4 axles artic	1.6056	1.5
Buses & coaches (million veh-km)	5184.5	5200
HGVs: 5 axles artic	6.5187	6
HGVs: 6 or more axles artic	6.004	5.7
Cars & taxis (billion veh-km)	401.75	420

e.g. I want my table to show relative amounts of mileage travelled by different vehicles in a year.

Audience: summary report to Ministry of Transport
(published for general public)

Table 1b: Motorised traffic levels by vehicle type (2008)

Vehicle type	Traffic (billion veh-km)
Cars & taxis	402
LGVs	68
HGVs	29
Buses & coaches	5
Motorcycles	5
All vehicles	509

1 Data aim

2 When to use tables and graphs

3 Principles for tables

4 Principles for graphs

5 What graph for my data?

3: Principles for tables

- *Always use **consistent units***
- **Round data consistently** in summary tables
- Use **captions and row and column headings** to show what the units are and what the numbers mean
- Right justify numbers in columns (or at least make all units, tens, hundreds etc. line up)
- Make sure that all tables in a publication are in a similar format
- Show time either from left to right or top to bottom
- Show row totals to the right and column totals at the bottom
- Put data to be compared in columns rather than rows
- *Keep tables as simple and clear as possible*

*Always show the **sample sizes** on which percentages are based*

- Make row percentages add up to 100
- Make column percentages add up to 100
- Make 100 across whole table

Indicate if values do not sum to 100% due to rounding

Percentages in tables - 1

Table 2: Number of traffic offences spotted by Police Officers, by road and type of offence.

Road	Speed	Mobile phone	Seatbelt	No Insurance
Road A	56	127	211	9
Road B	182	126	218	10
Road C	123	179	236	9
Road D	81	76	215	8

Table 3: Distribution of each type of traffic offence spotted by Police Officers, by road and type of offence

Road	Speed	Mobile phone	Seatbelt	No Insurance
Road A	13%	25%	24%	25%
Road B	41%	25%	25%	28%
Road C	28%	35%	27%	25%
Road D	18%	15%	24%	22%
Sample size	442	508	880	36

Percentages in tables - 2

Table 2: Number of traffic offences spotted
by Police Officers, by road and type of offence.

Road	Speed	Mobile phone	Seatbelt	No Insurance
Road A	56	127	211	9
Road B	182	126	218	10
Road C	123	179	236	9
Road D	81	76	215	8

Table 4: Distribution of traffic offences
spotted by Police Officers across each road

Road	Speed	Mobile phone	Seatbelt	No Insurance	Total
Road A	14%	32%	52%	2%	403
Road B	34%	24%	41%	2%	536
Road C	22%	33%	43%	2%	547
Road D	21%	20%	57%	2%	380

Percentages in tables - 3

**Table 2: Number of traffic offences spotted
by Police Officers, by road and type of offence.**

Road	Speed	Mobile phone	Seatbelt	No Insurance
Road A	56	127	211	9
Road B	182	126	218	10
Road C	123	179	236	9
Road D	81	76	215	8

**Table 5: Distribution of different types of traffic offences and roads
(sample size = 1866)**

Road	Speed	Mobile phone	Seatbelt	No Insurance
Road A	3%	7%	11%	<1%
Road B	10%	7%	12%	1%
Road C	7%	10%	13%	<1%
Road D	4%	4%	12%	<1%

1 Data aim

2 When to use tables and graphs

3 Principles for tables

4 Principles for graphs

5 What graph for my data?

Understand:

- The type of data to be presented
- The key feature to be portrayed
- How the information will be used
- The intended audience

Figure 2a: Hardware parts market profit per part (\$k)

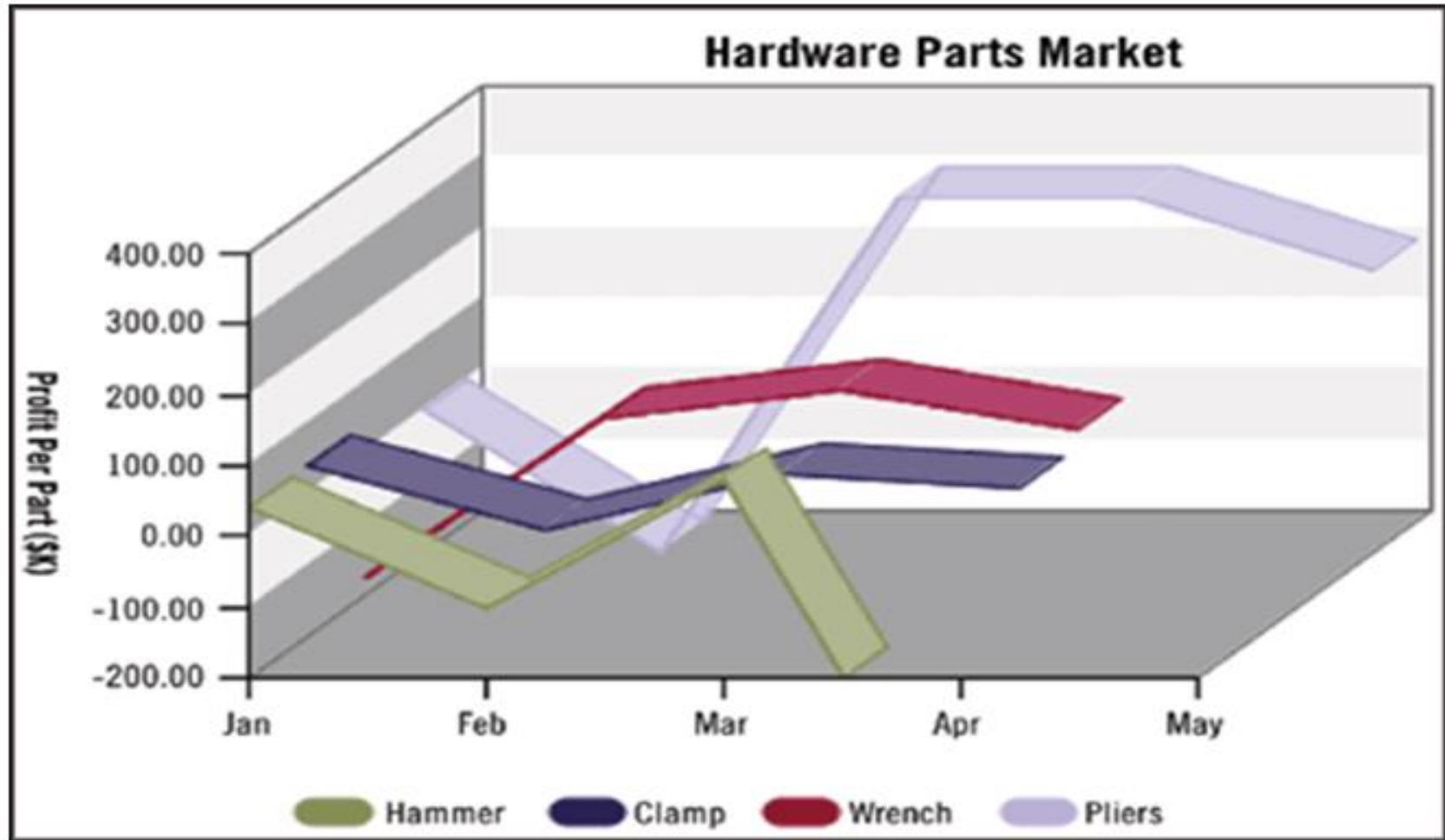
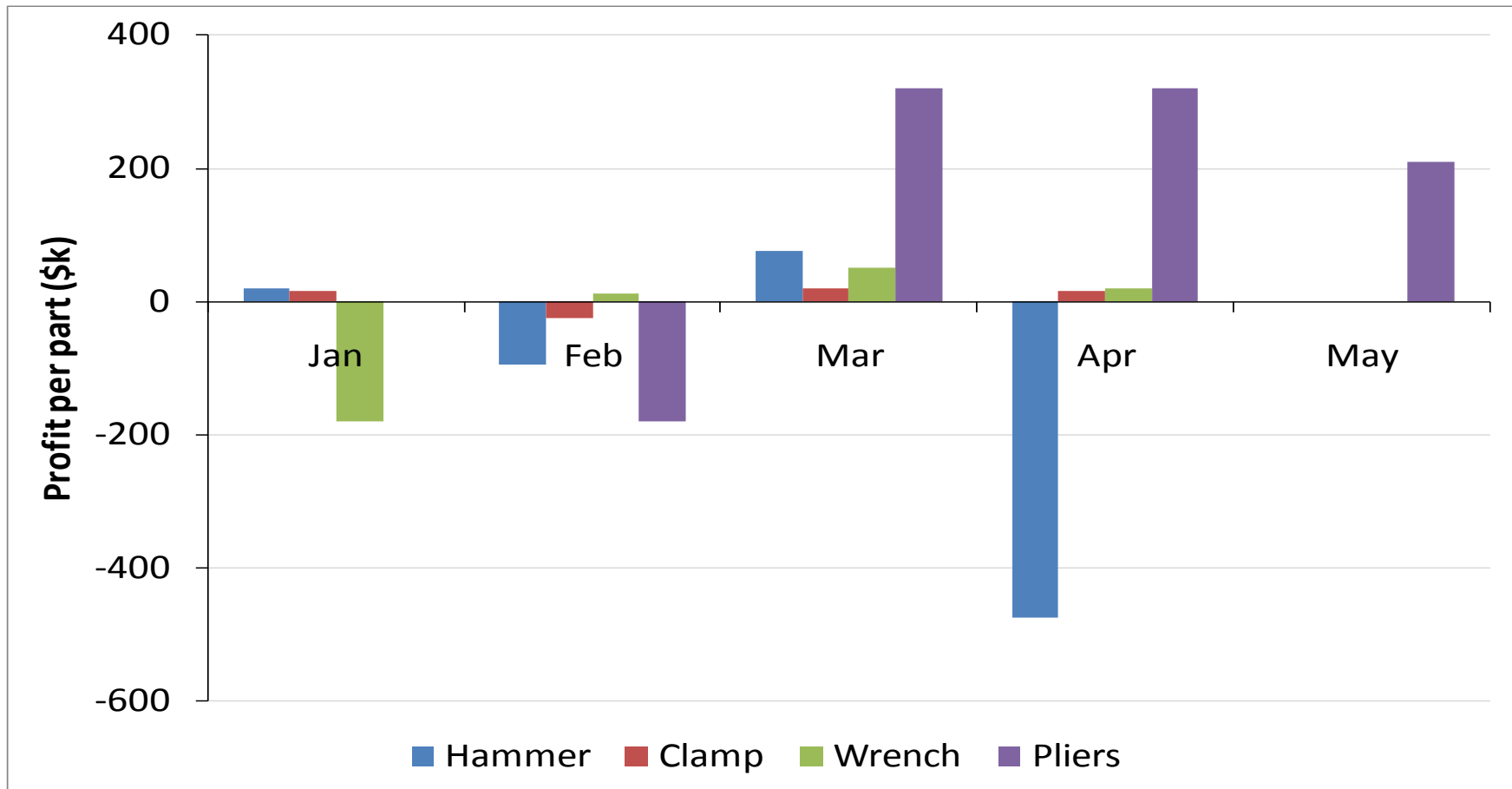
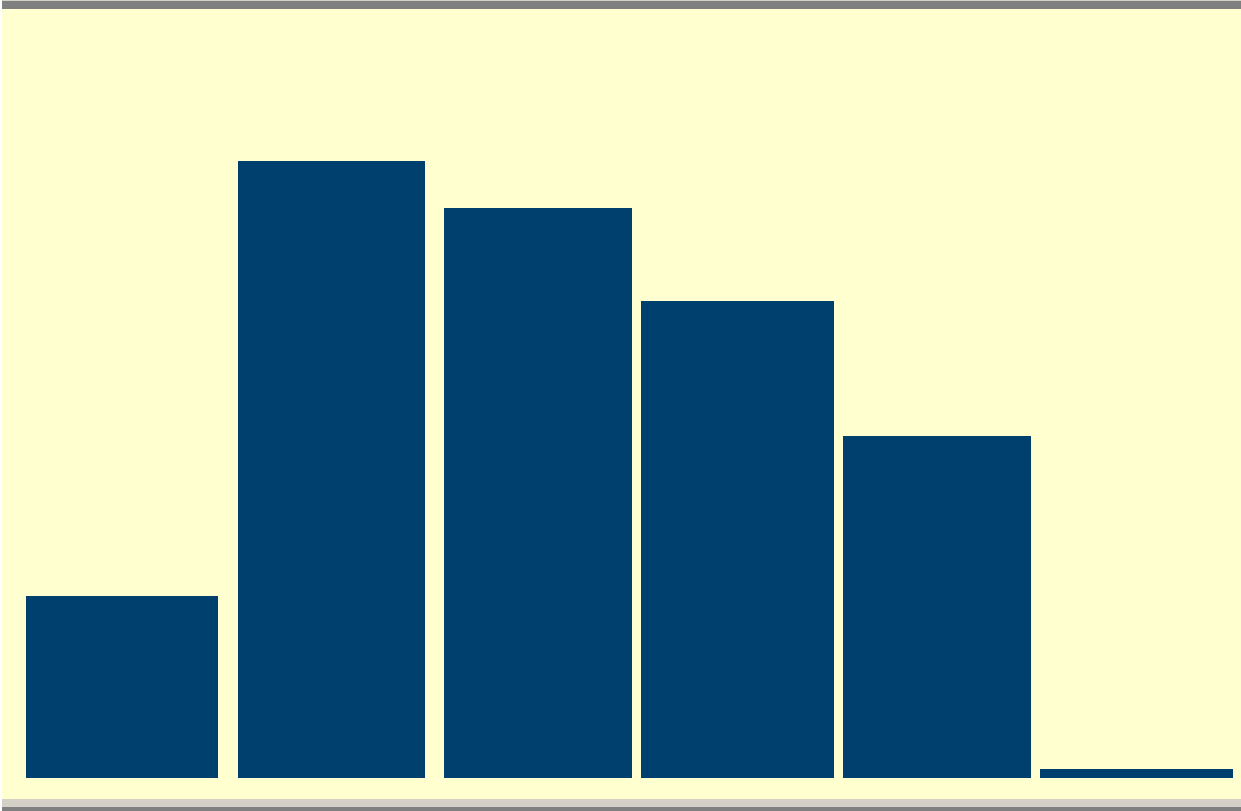


Figure 2b: Hardware parts market profit per part (\$k)

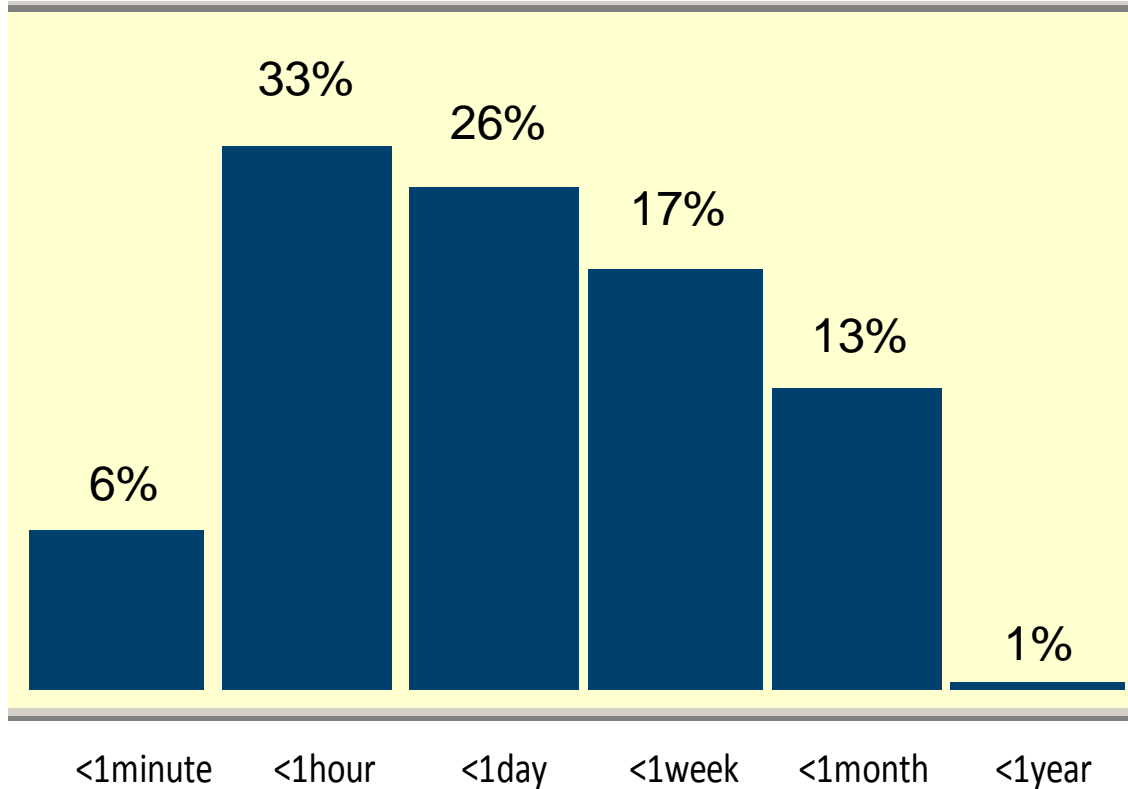


*IT Call duration,
last 6 months*

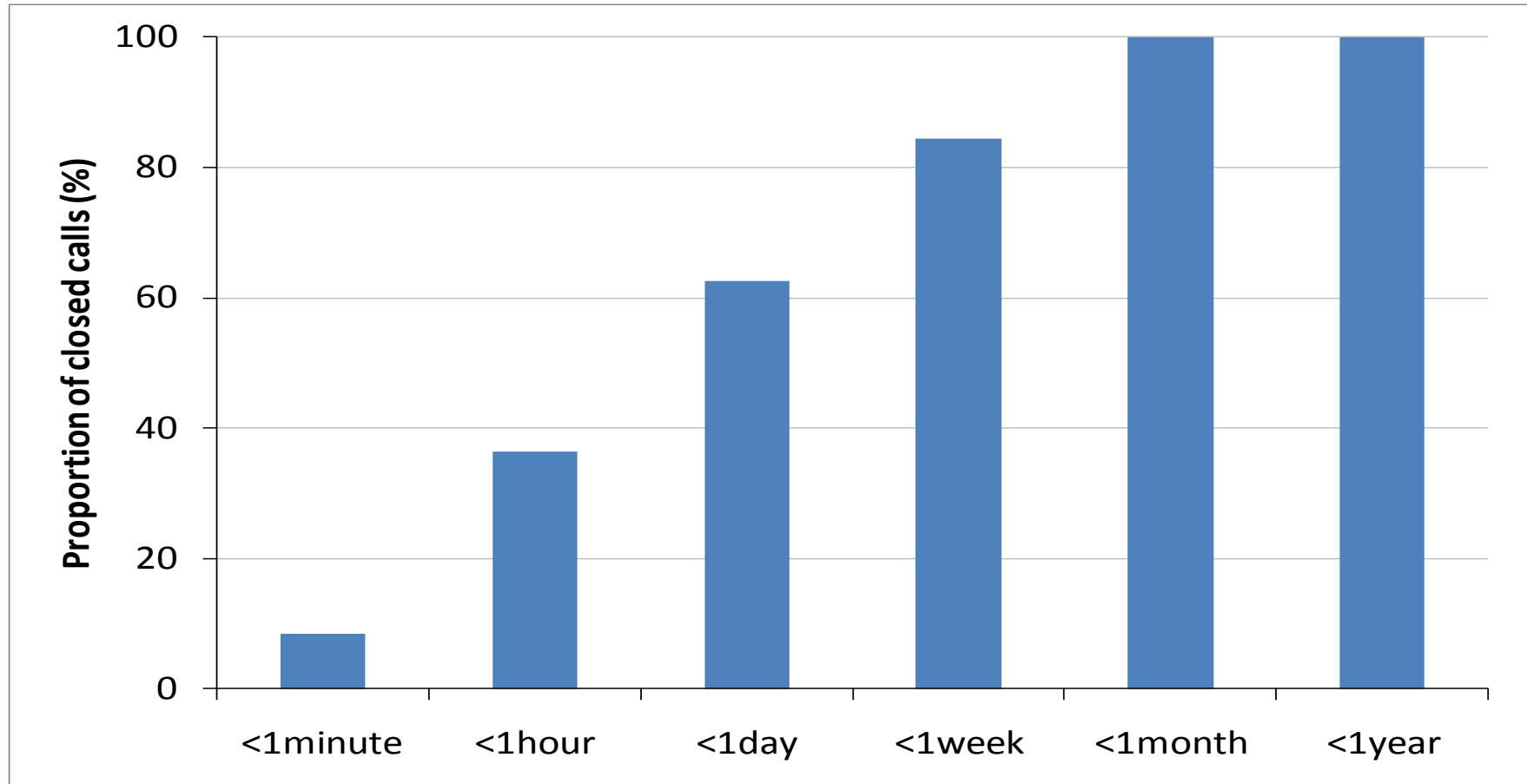


Example: IT enquiries – 2 An improvement?

*IT Call duration,
last 6 months*

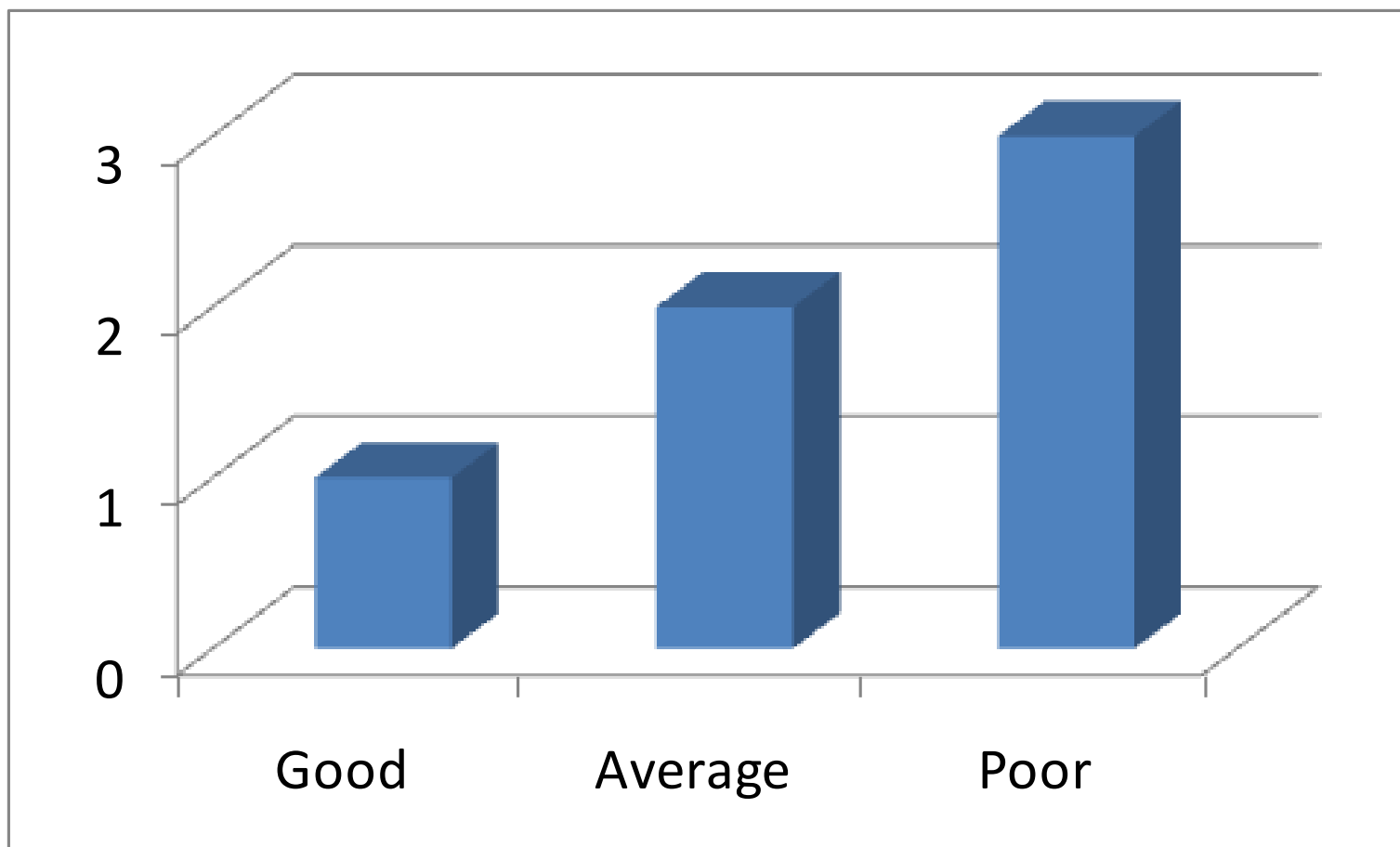


of duration of closed calls over 6 months



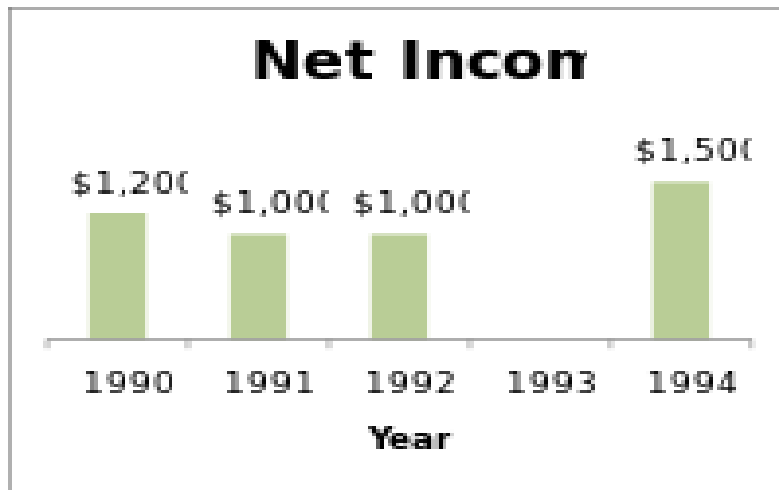
Example: 3D charts

DO NOT USE THEM!

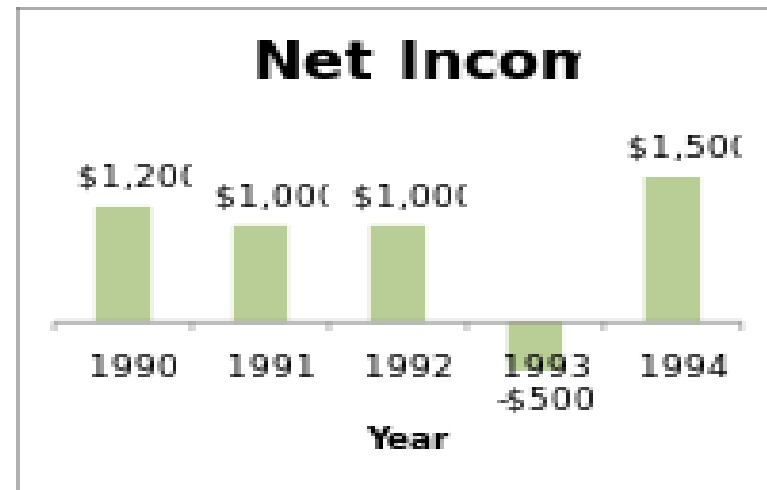


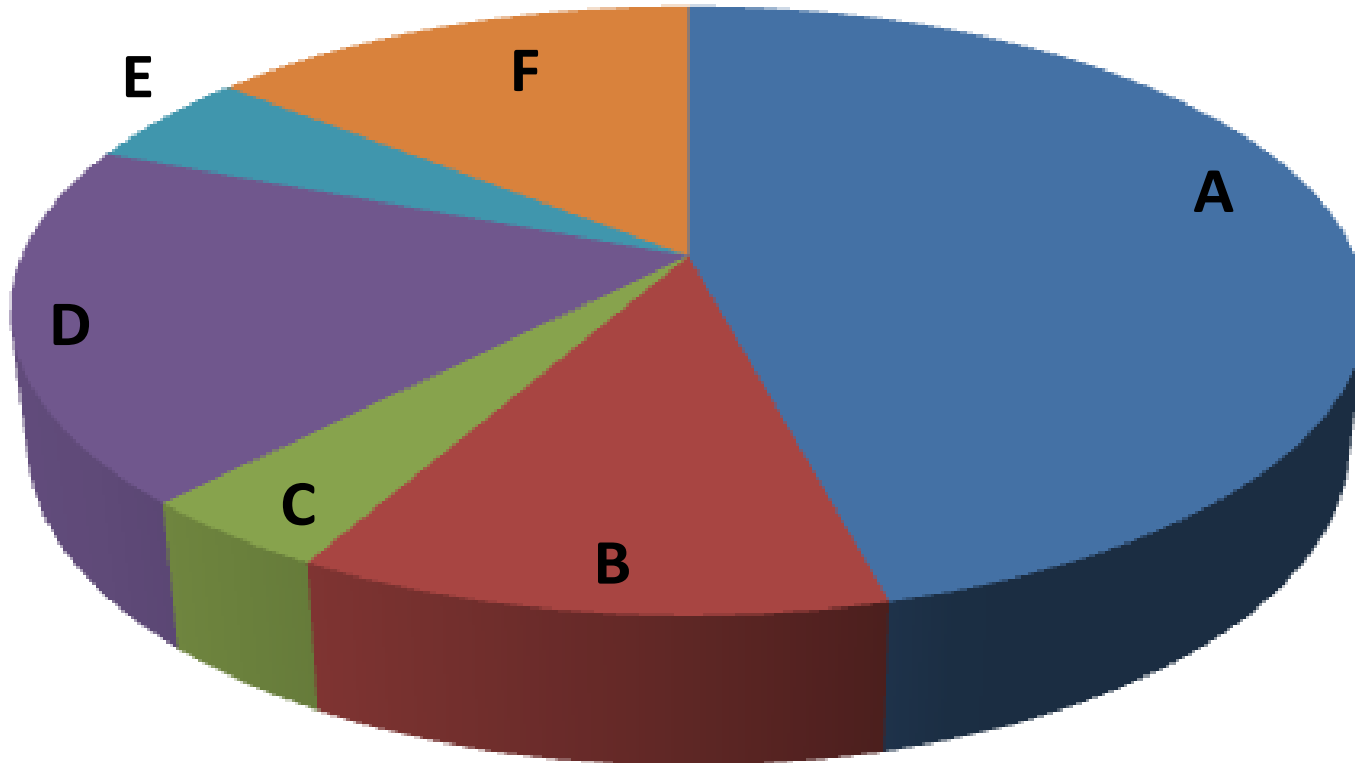
Example: Bar graph omitting negative value

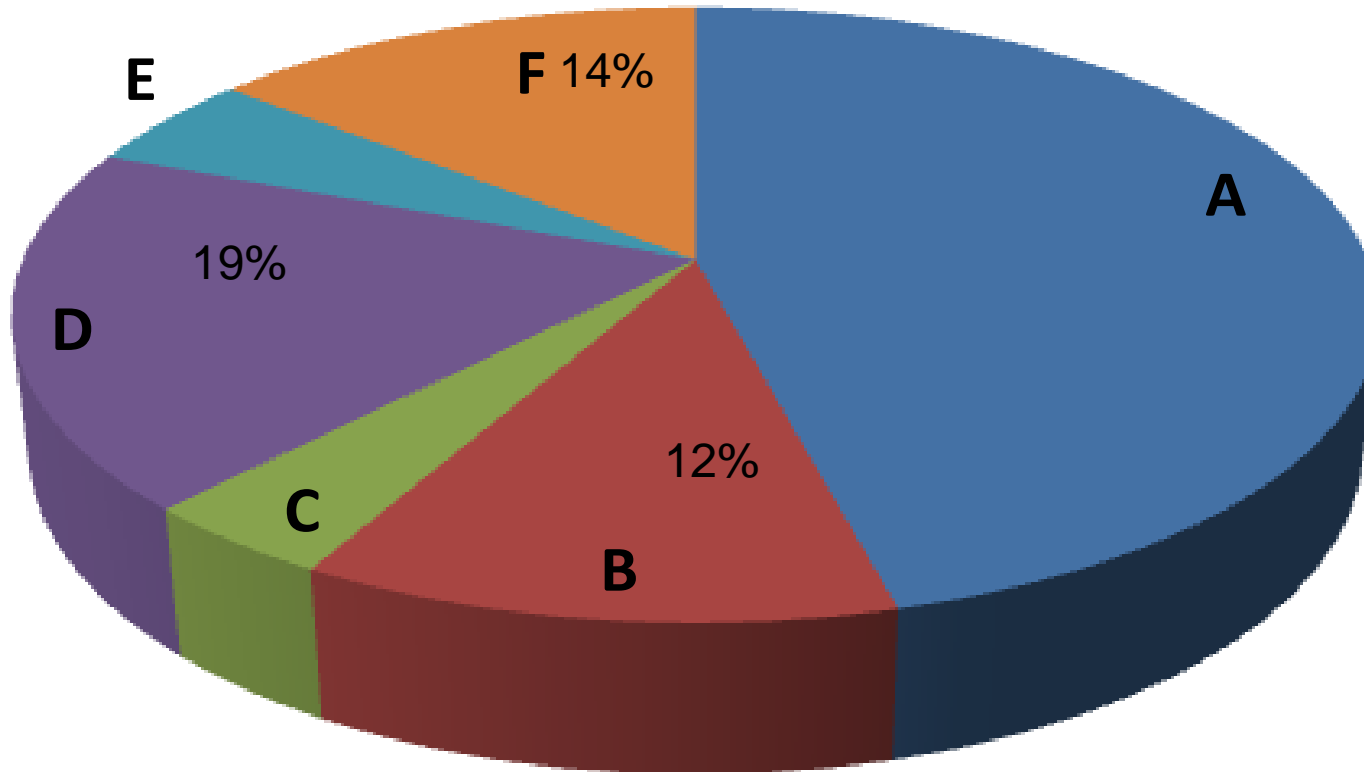
Graph omitting negative value



Regular graph



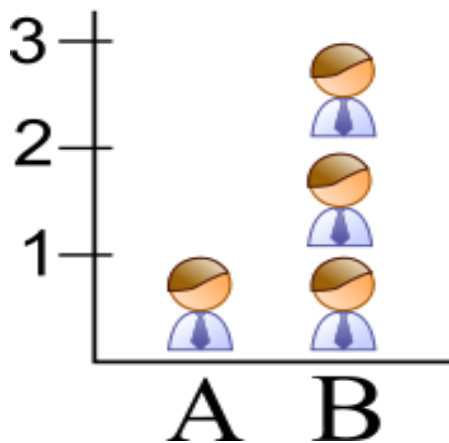




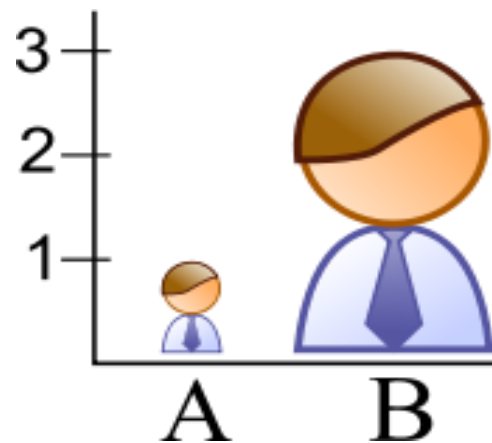
Example: Scaling

Improper pictogram scaling in bar graph

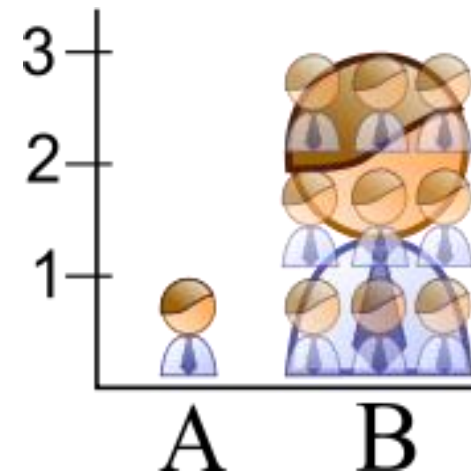
Comparison



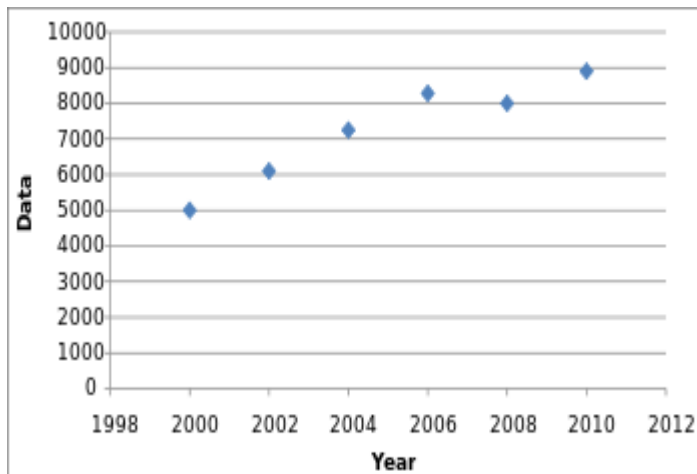
Improper scaling



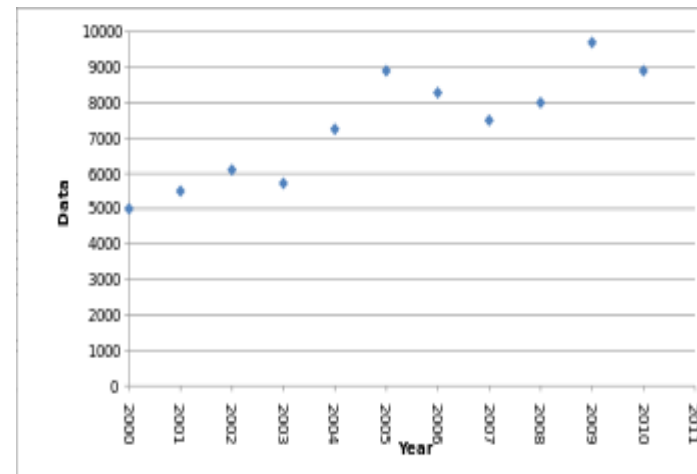
What it hides



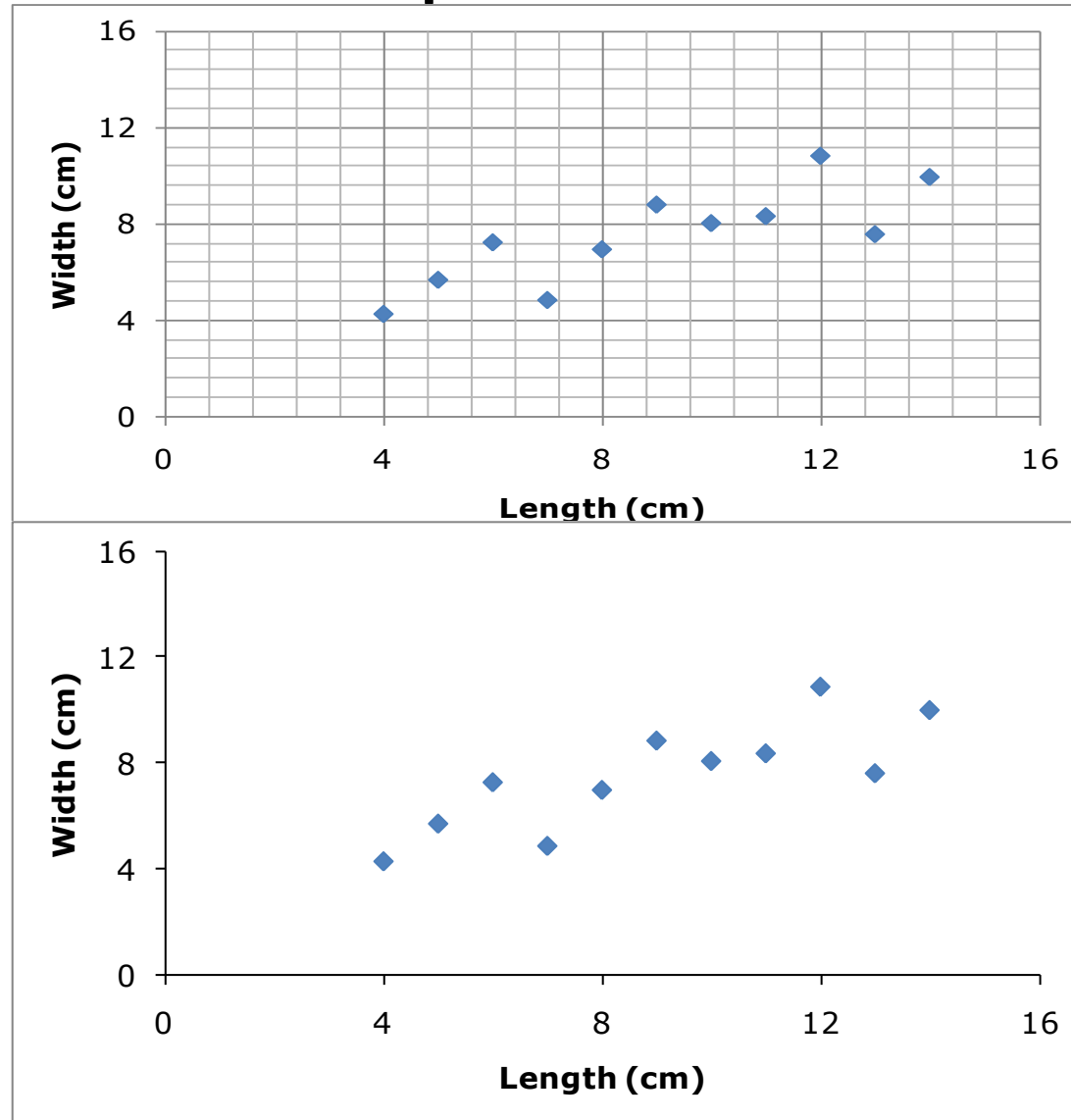
Scatter plot with missing categories

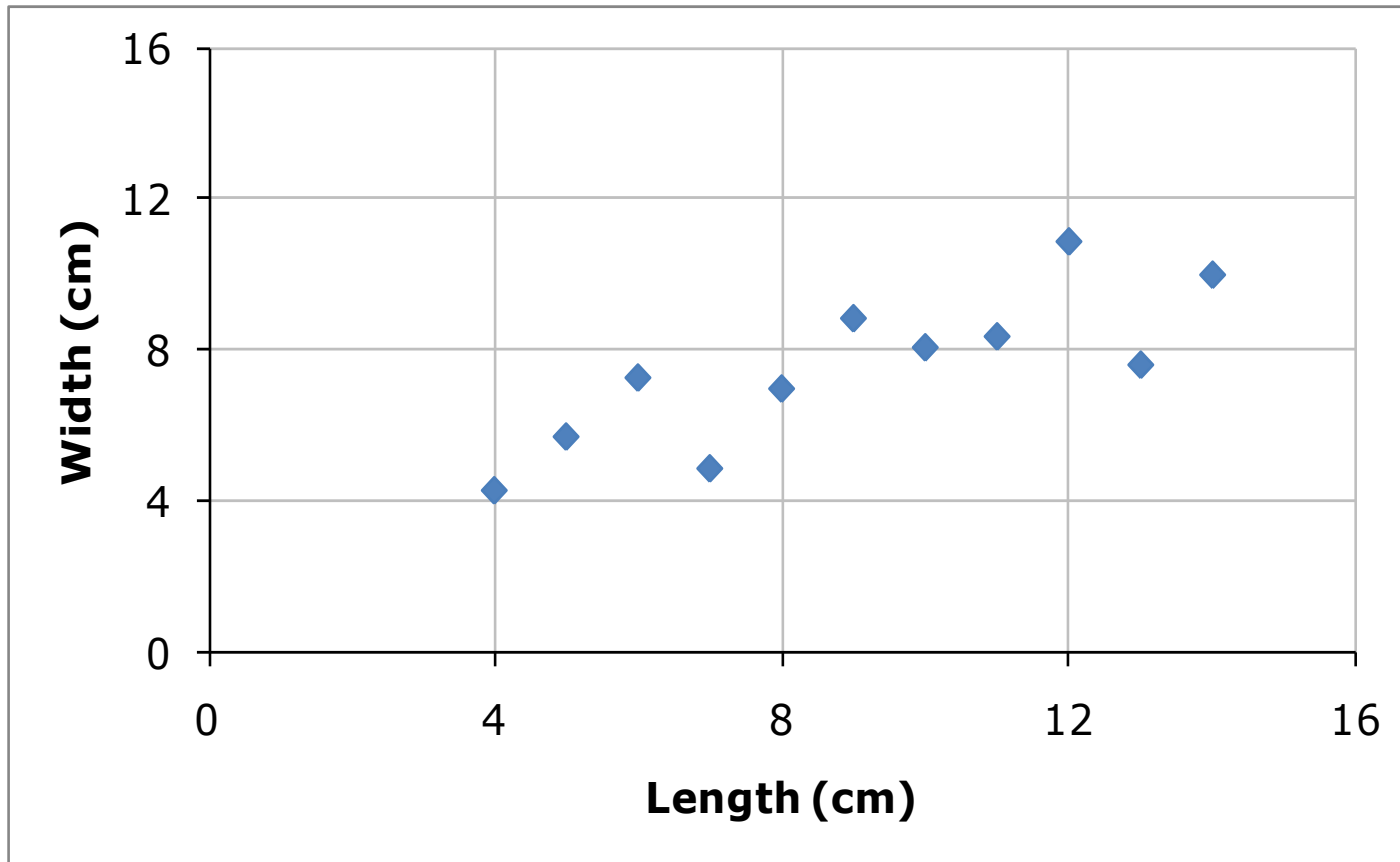


Regular scatter plot



Example: Gridlines





Axes

- Axes can easily mislead – the scale is important
- Need “Z” (zigzag) if broken
- Axes should be heavier than gridlines

AFCAP

Example:
Axes

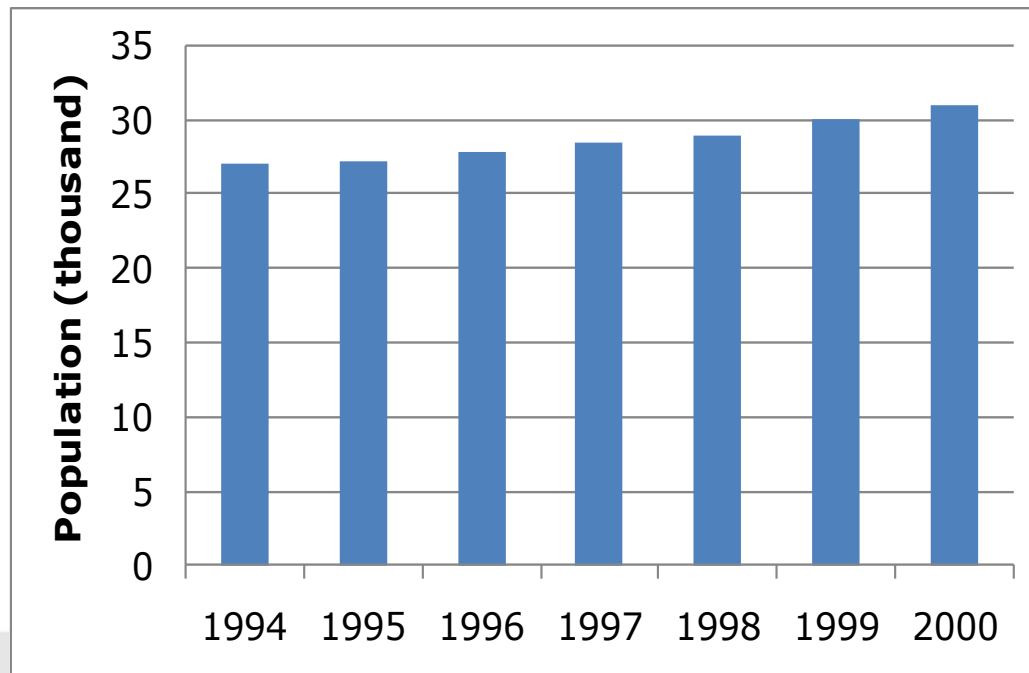
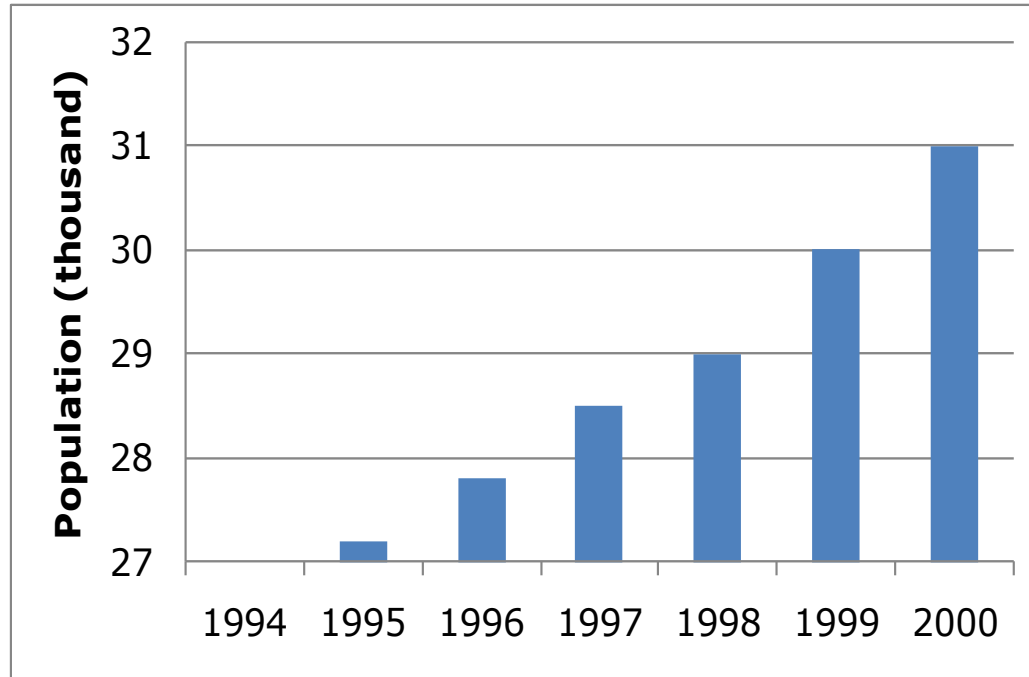


Table 6: Population (real in thousands and indexed to 1994) of a town and a city in Ethiopia from 1994 to 2000.

Year	Town		City	
	Population	Index	Population	Index
1994	27.0	100	562	100
1995	27.2	101	575	102
1996	27.8	103	590	105
1997	28.5	106	625	111
1998	29.0	107	650	116
1999	30.0	111	700	125
2000	31.0	115	800	142

Figure: Relative population figures of a town and a city indexed to 1994 values

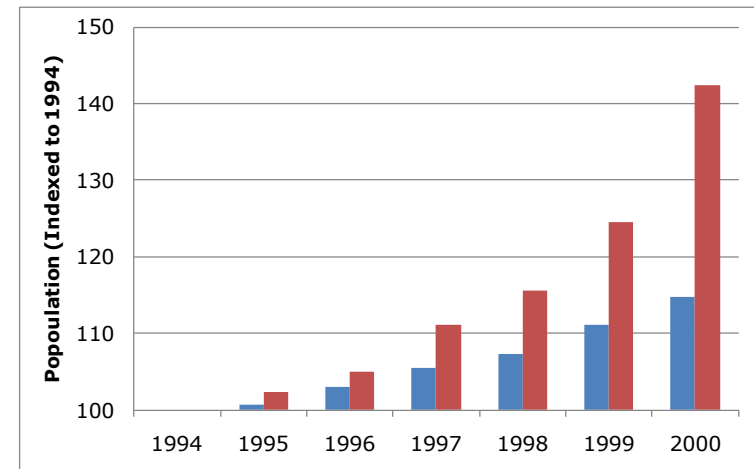


Figure 10: Plot of blood concentration against saliva concentration for 26 subjects, by gender

Have you labelled...

1. your graph?
2. your axes
(including units)?
3. your series?

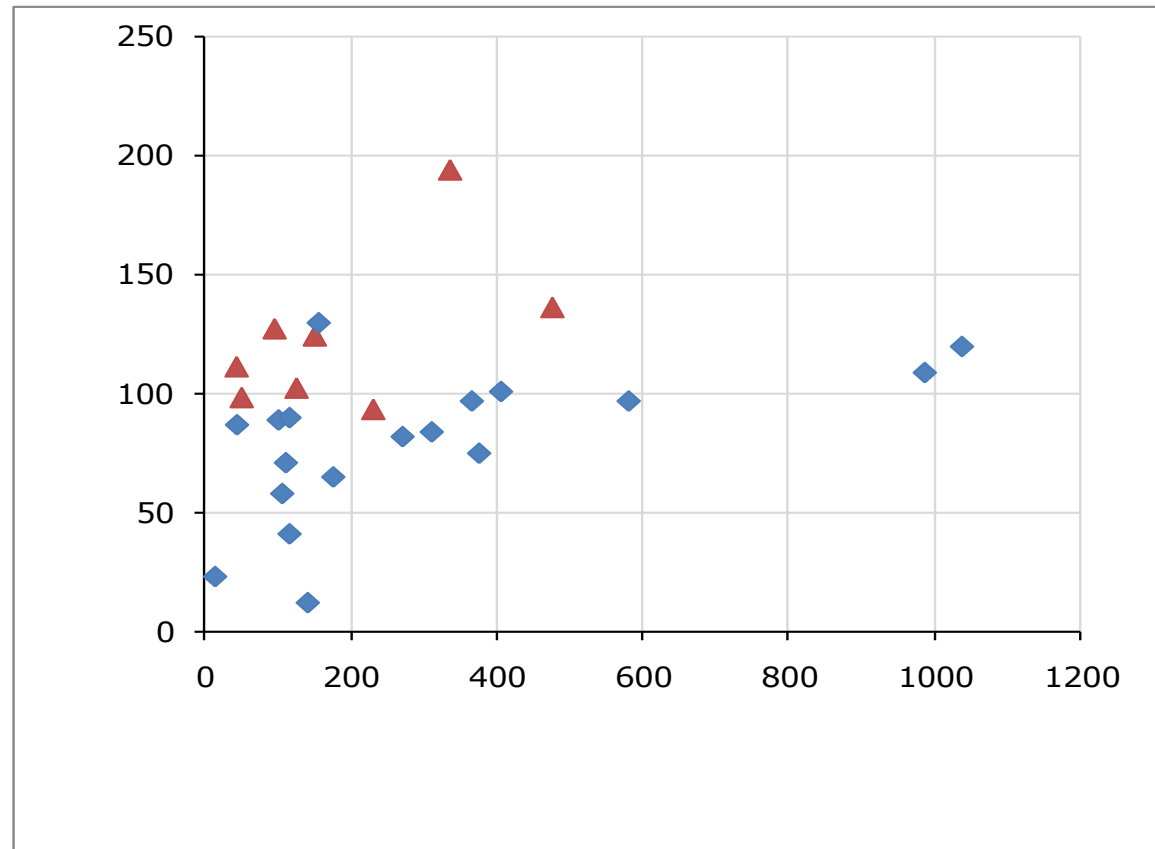


Figure 10: Plot of blood concentration against saliva concentration for 26 subjects, by gender

Have you labelled...

1. your graph?
2. your axes
(including units)?
3. your series?

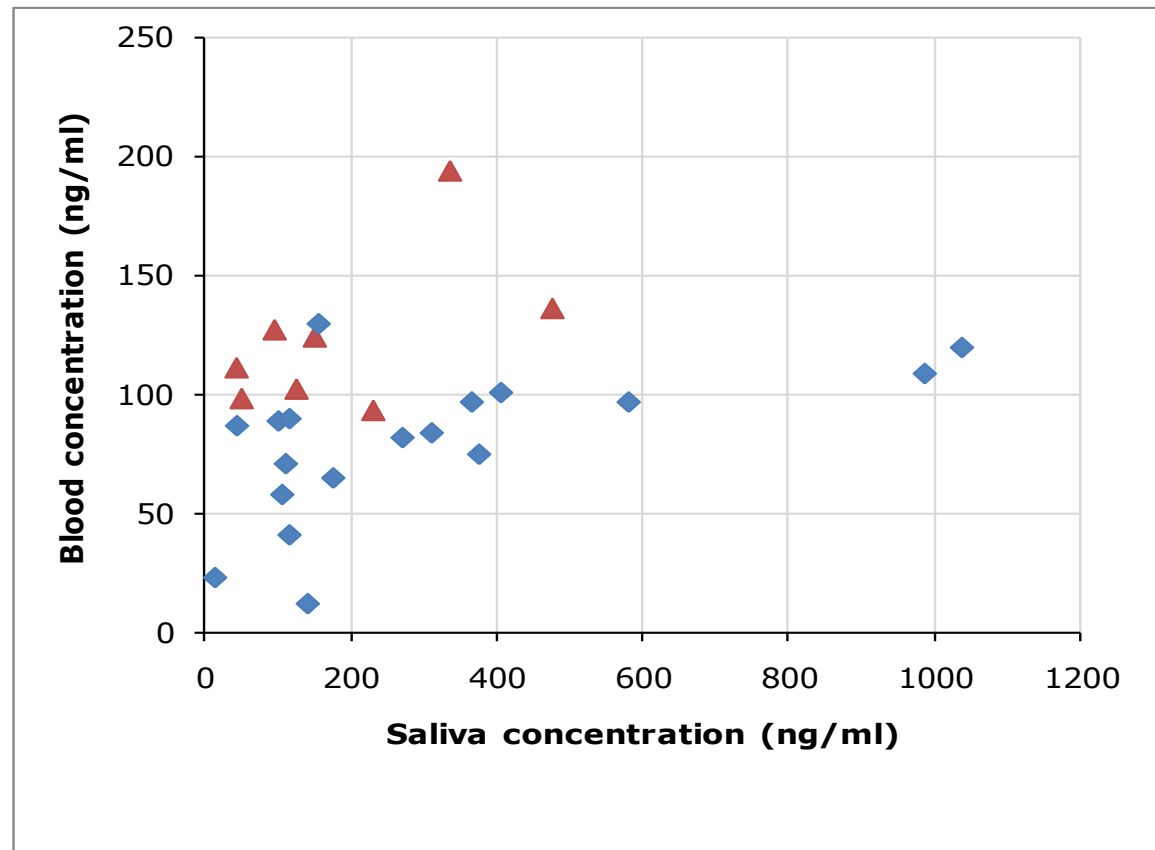
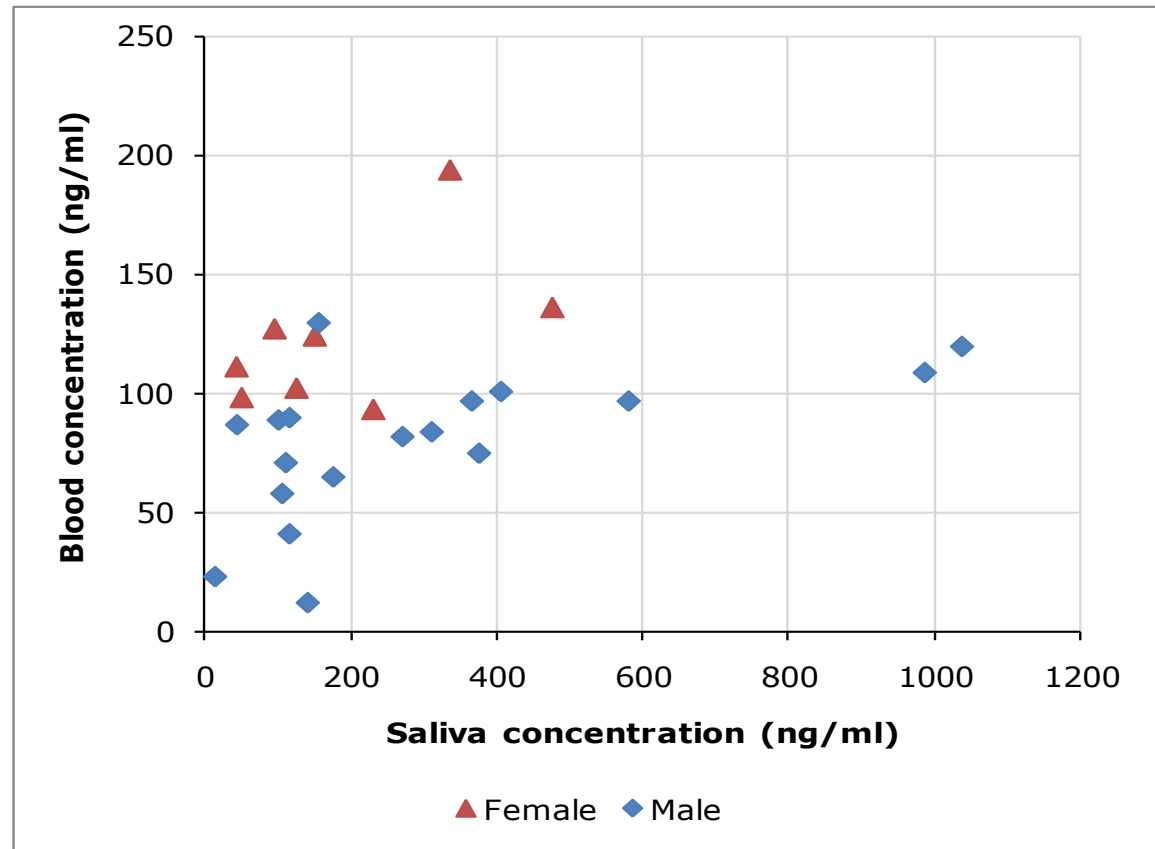


Figure 10: Plot of blood concentration against saliva concentration for 26 subjects, by gender

Have you labelled...

1. your graph?
2. your axes
(including units)?
3. your series?



1 Data aim

2 When to use tables and graphs

3 Principles for tables

4 Principles for graphs

5 What graph for my data?

4: Principles for graphs

- *Never use 3D graphs (unless you have 3D data)*
- Use colour sparingly – shades of a single colour are often more effective
- *Always put the zero point on the scale when using absolute numbers*
- *Label diagrams clearly*
- Show gridlines to aid interpretation
- Round numbers on the axes appropriately
- *Ensure the picture accurately represents the data*
- Make the diagram simple enough so that the reader can quickly assimilate the message

Discrete, grouped, nominal or ordinal data

Examples of data type:

- Type of motorbikes ridden
- Accident involvement
- Age and gender

Figure 4: Number of fatal accidents on different road types in 2008

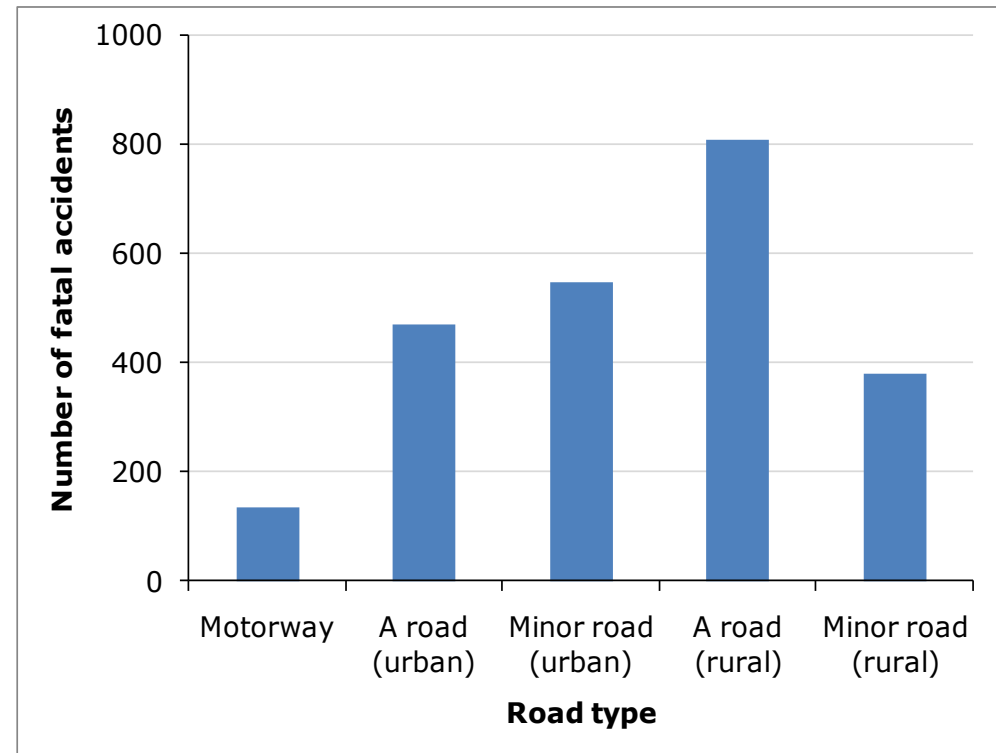


Figure 5: Type of bike ridden in the last 12 months by gender of rider (in response to bike survey)

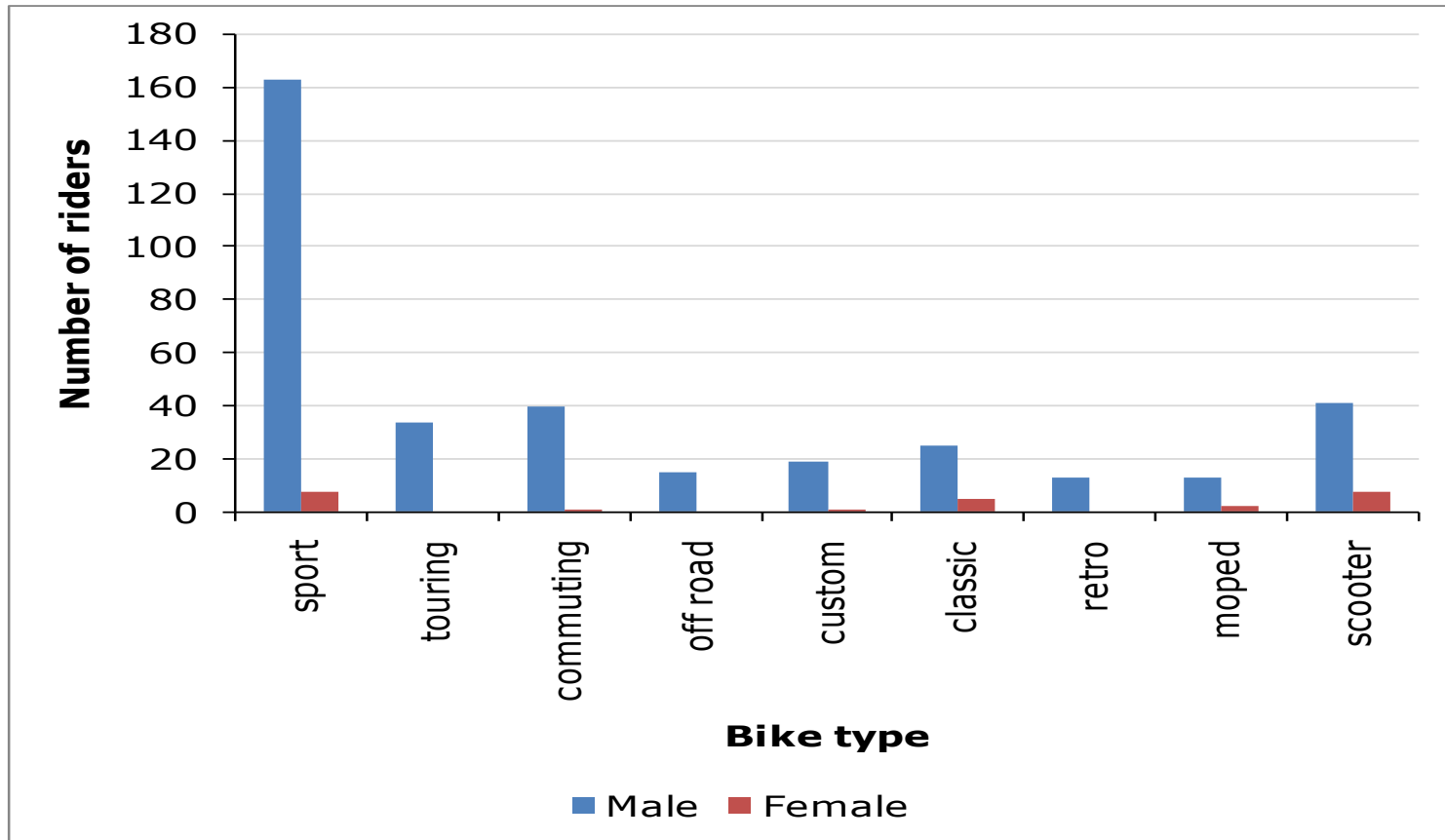
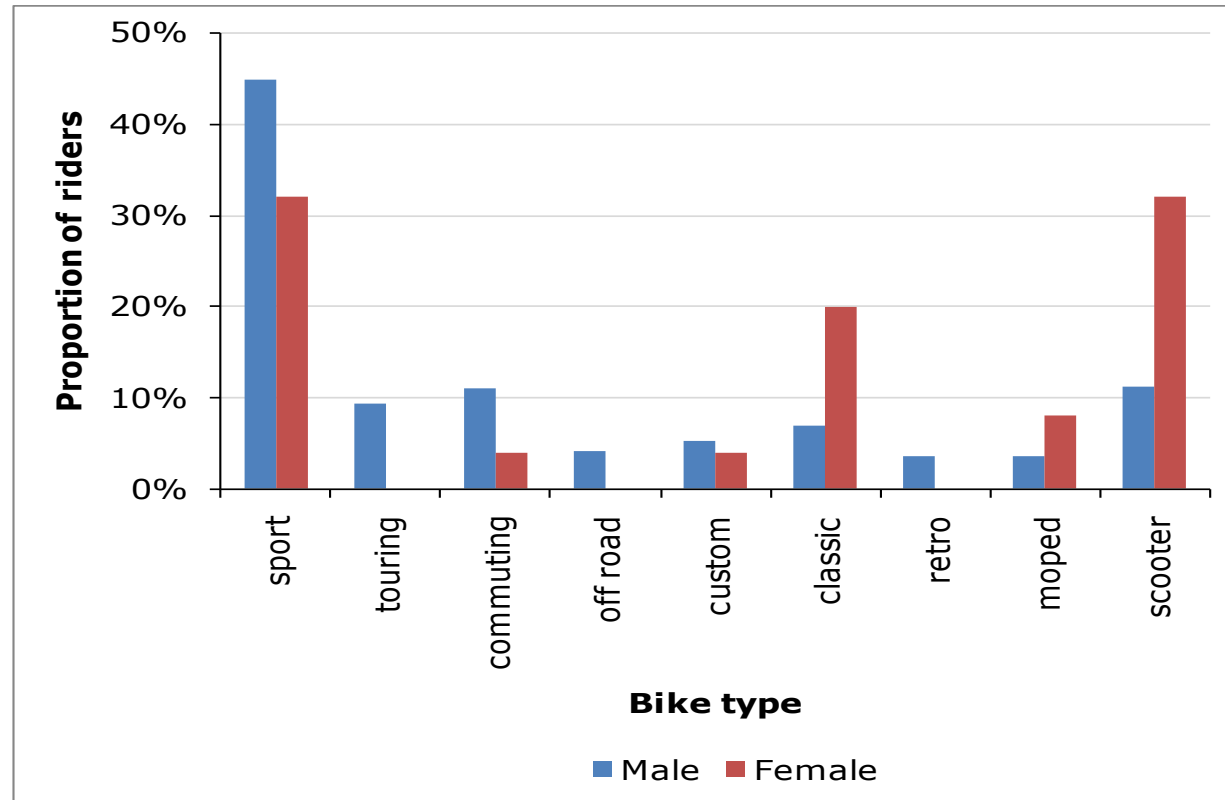


Figure 6: Type of bike ridden in the last 12 months by gender of rider (response to bike survey). 365 male responders, 25 female responders

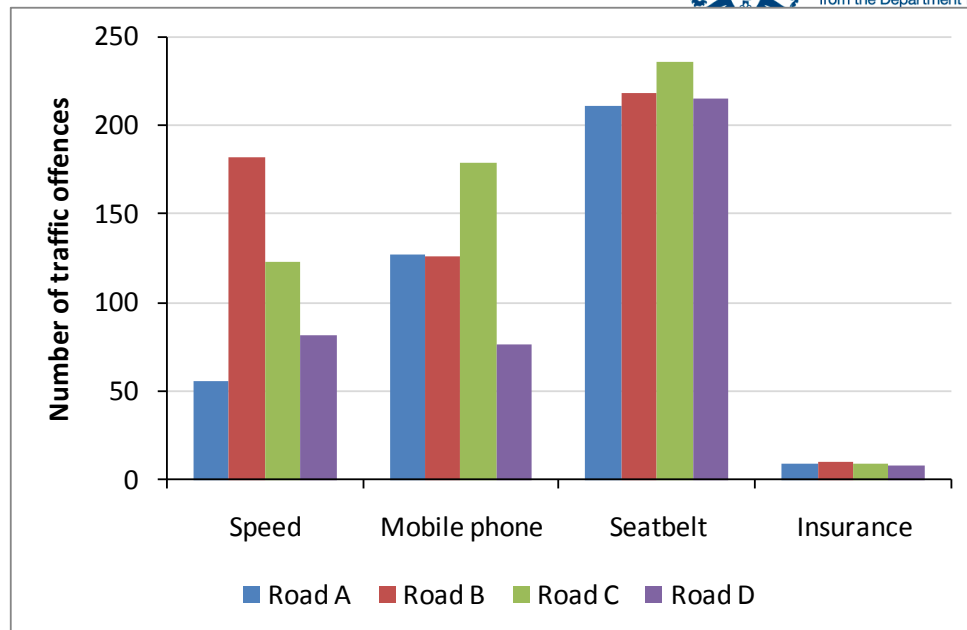


Discrete, grouped, nominal or ordinal data

- 1. If there is no obvious ordering, order bars by size (either increasing or decreasing)*
- 2. Include the zero point on a scale, otherwise 40 and 20, for example, are not proportionate*
3. When comparing more than one variable, group the bars to be compared
- 4. Don't overlap bars – the visible area is smaller for overlapped bars*
5. Stacked bar charts are good at displaying variation in the first category and the total but not other categories



For comparing differences across roads.



For comparing differences across traffic offence type.

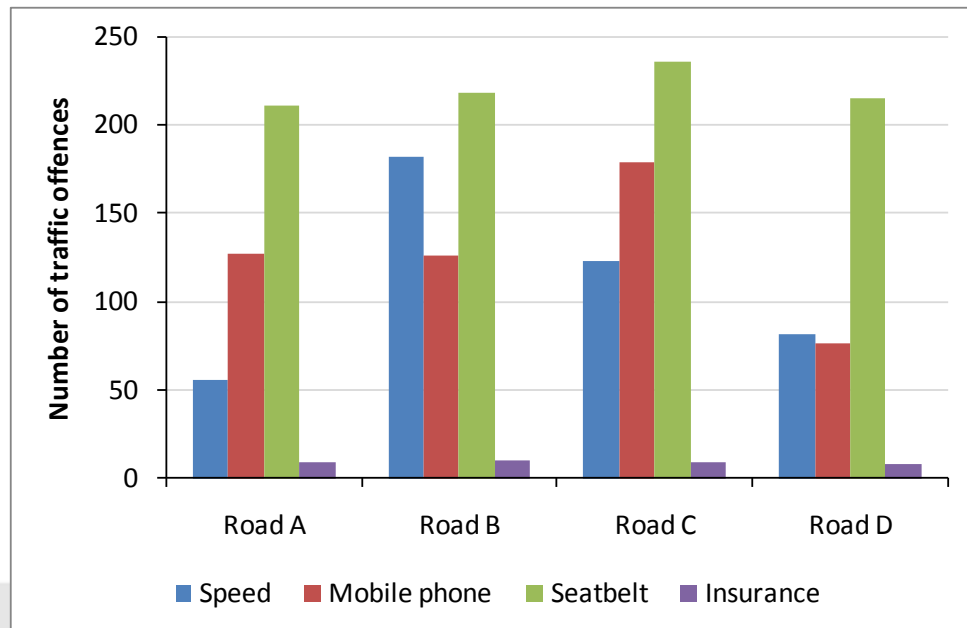


Figure 4: Number of traffic offences observed on four roads in Arusha in one day

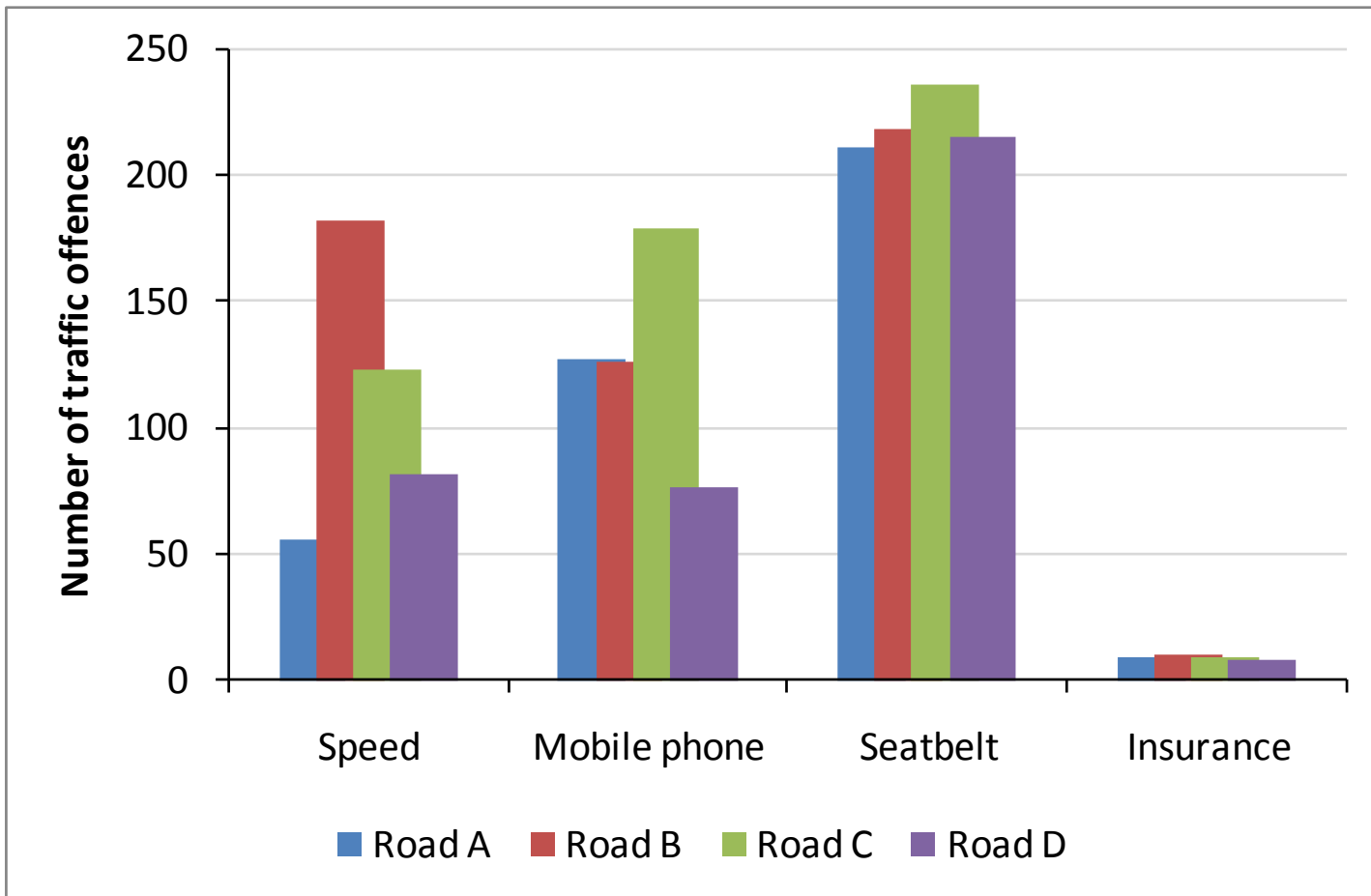


Figure 4: Number of traffic offences observed on four roads in Arusha in one day

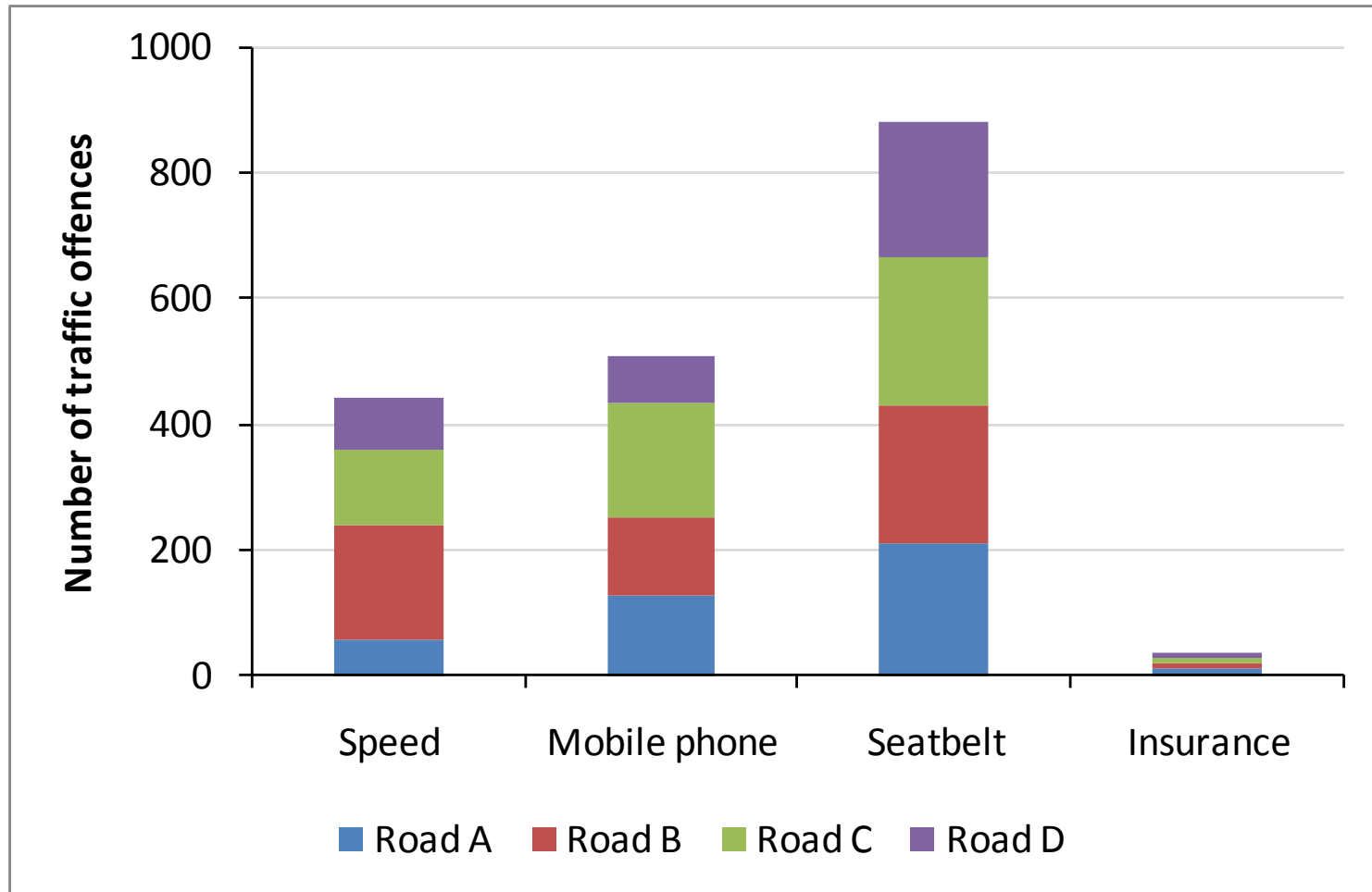
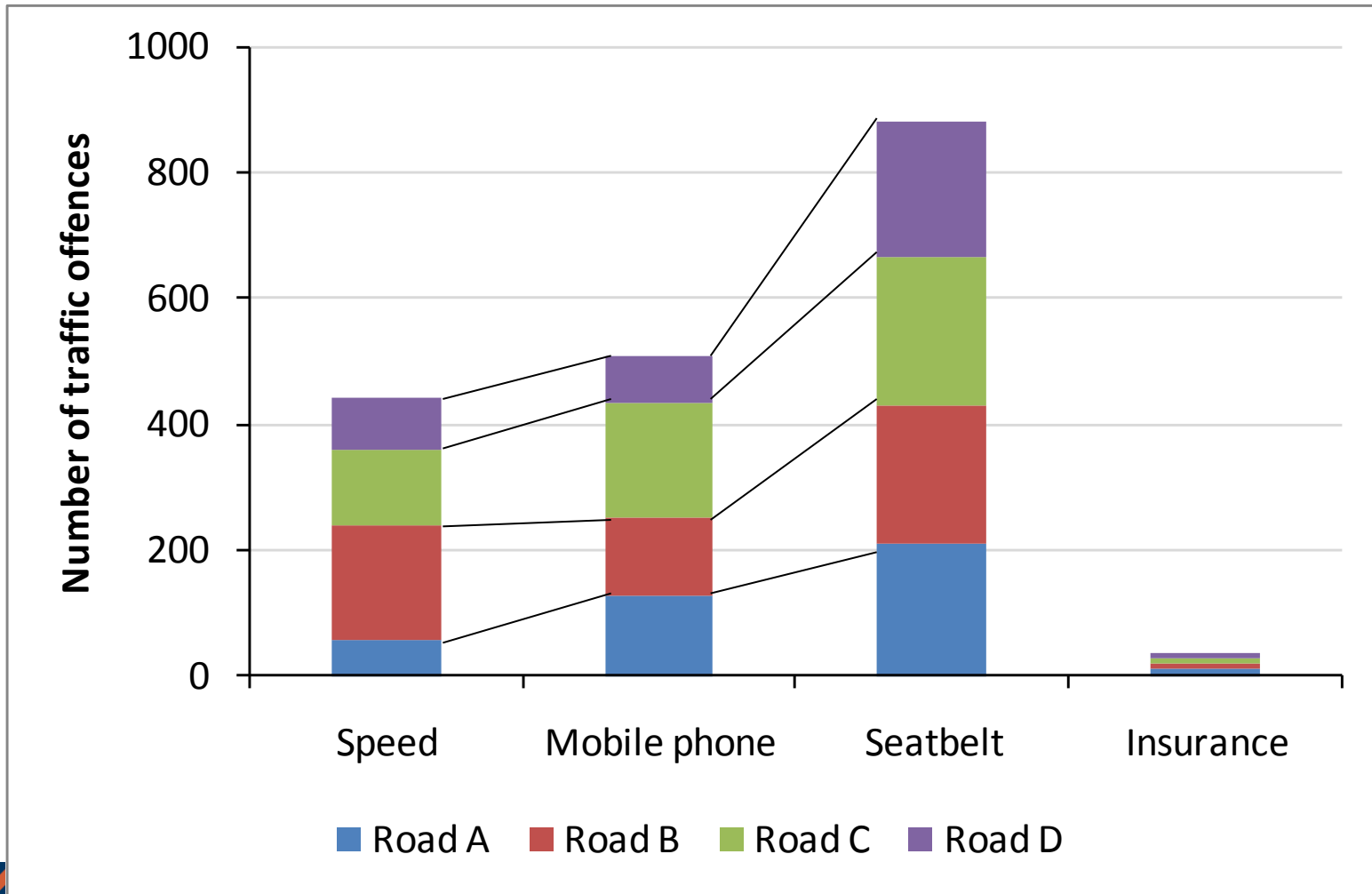


Figure 4: Number of traffic offences observed on four roads in Arusha in one day



1 Data aim

2 When to use tables and graphs

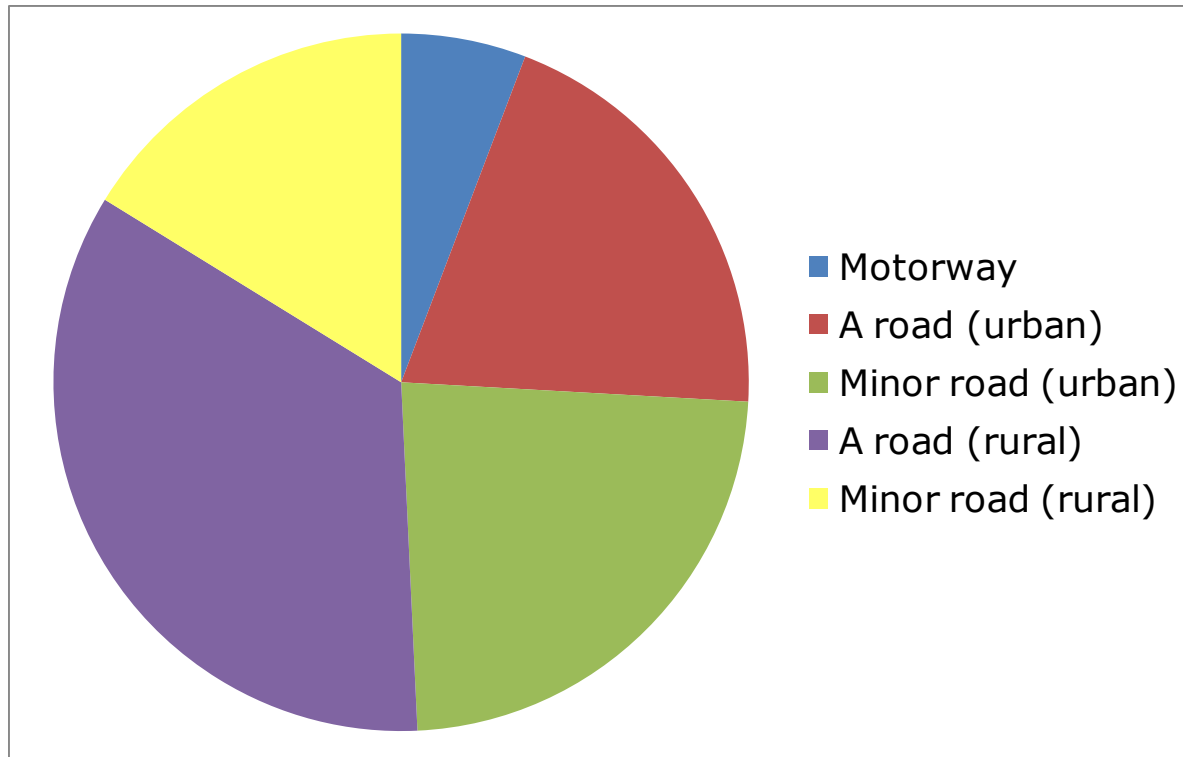
3 Principles for tables

4 Principles for graphs

5 What graph for my data?

Discrete, grouped, nominal data

Figure 7: The proportion of fatal accidents by road type in 2008



Discrete, grouped, nominal or ordinal data

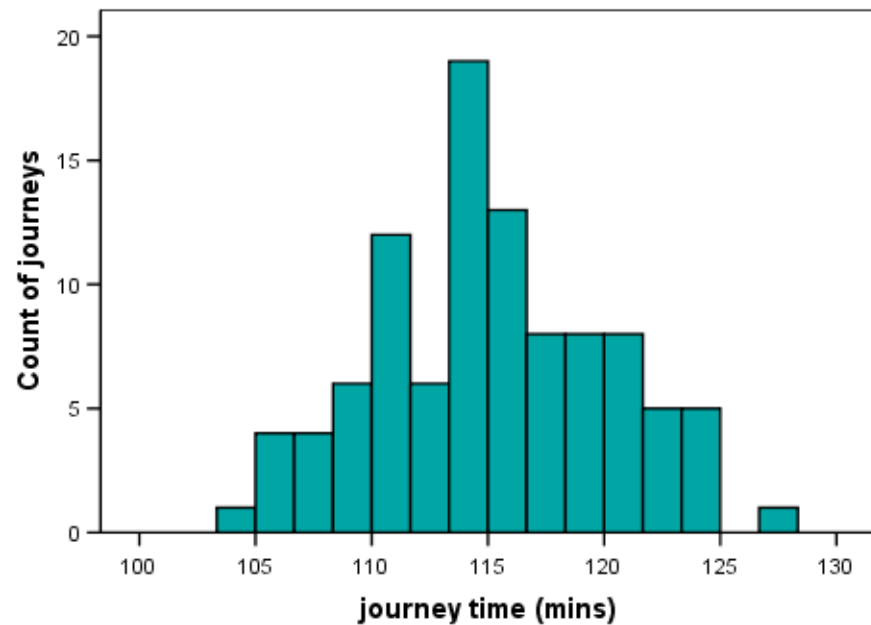
1. You have to have a really good reason for using a pie chart! Often a bar chart (or histogram) will present data in a more meaningful way.
2. *Pie charts are only useful for different sized segments.*
3. Pie charts should only be used when the total sum of all the segments has meaning.
4. Pie charts are effective for presenting a small number of pieces of data. *The optimal number of segments is 6* (between 3 and 10 is advised)

Continuous data

Figure 8: 100 journey times from Dar es Salaam to Morogoro

Examples:

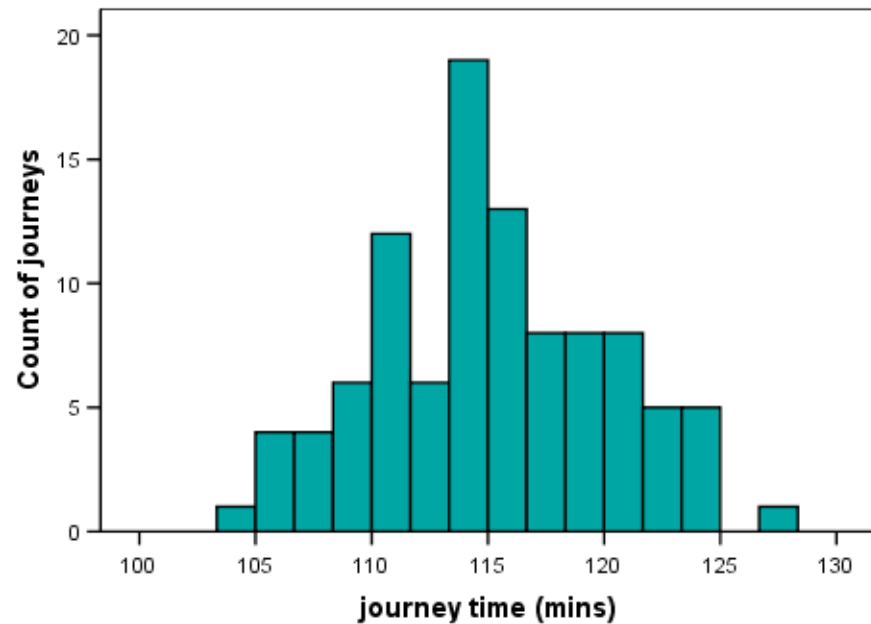
- Journey time from Dar es Salaam to Morogoro
- Length of motorway delays due to accident
- Measurements



Continuous data

1. Plot continuous data with histograms
2. Choose bar widths carefully
3. Start the Y-axis at 0 - or include \bar{N}
4. If absolute comparisons are necessary, start the X-axis at zero to avoid misinterpretation

Figure 8: 100 journey times from Dar es Salaam to Morogoro

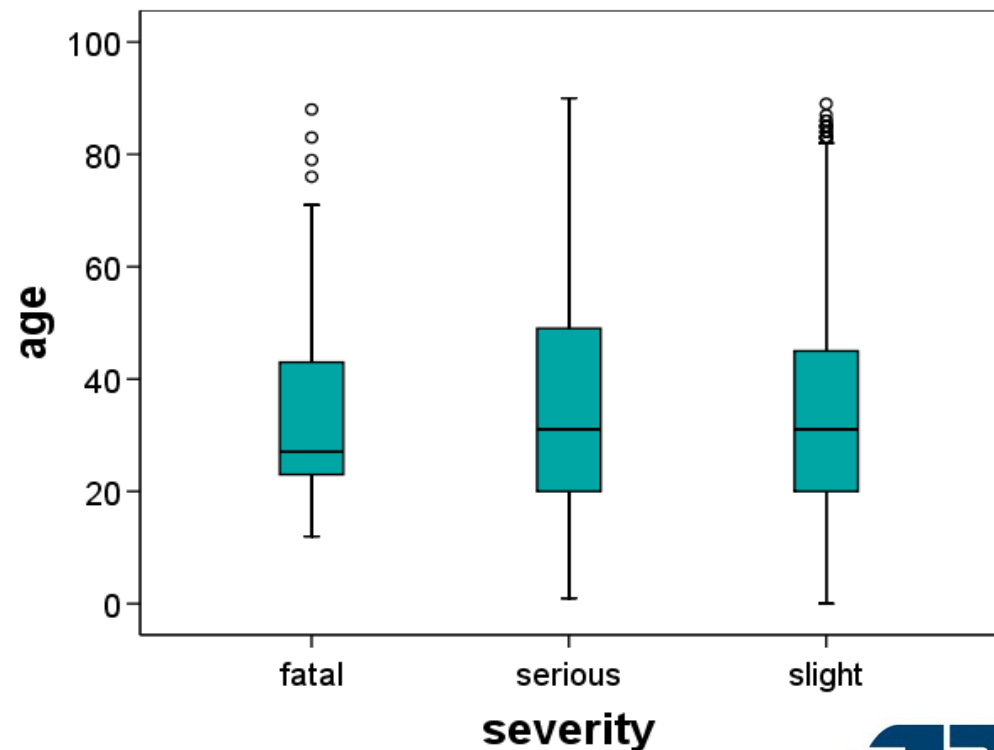


Continuous data (more than one group)

Examples:

- Heights of males and females
- Reaction times for group 1 and group 2
- Age distribution of casualties by severity

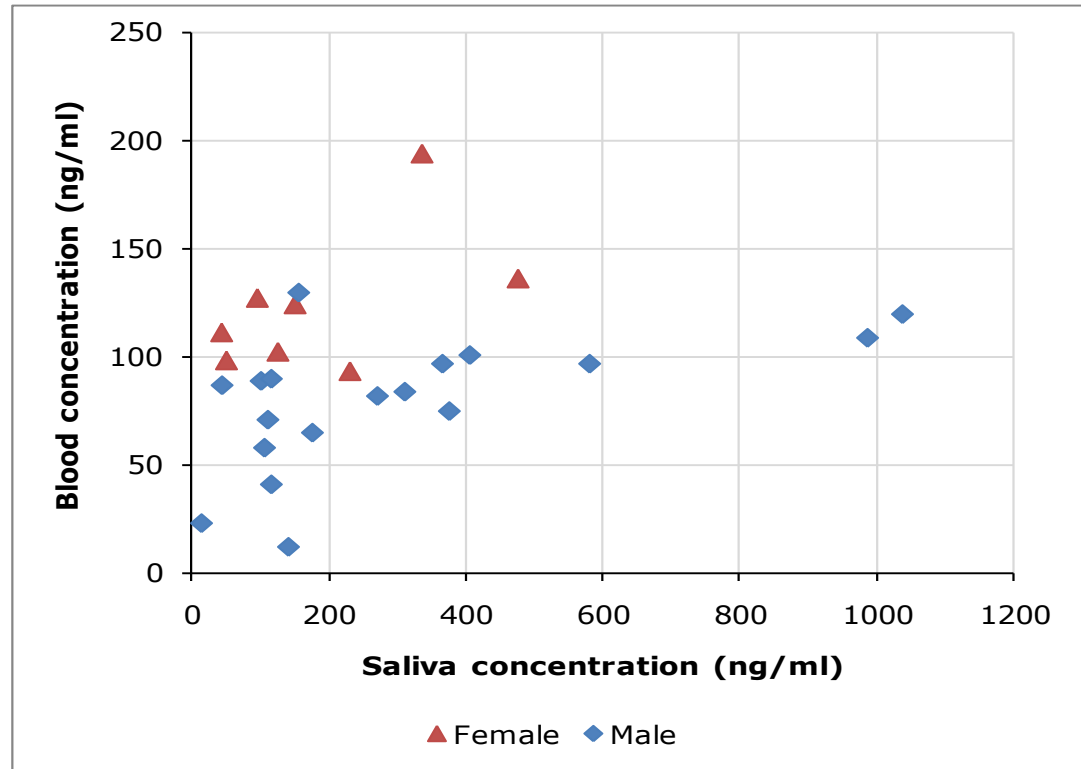
Figure 9: Age distributions of casualties in road accidents



Paired continuous data

Figure 9: Plot of blood concentration against saliva concentration

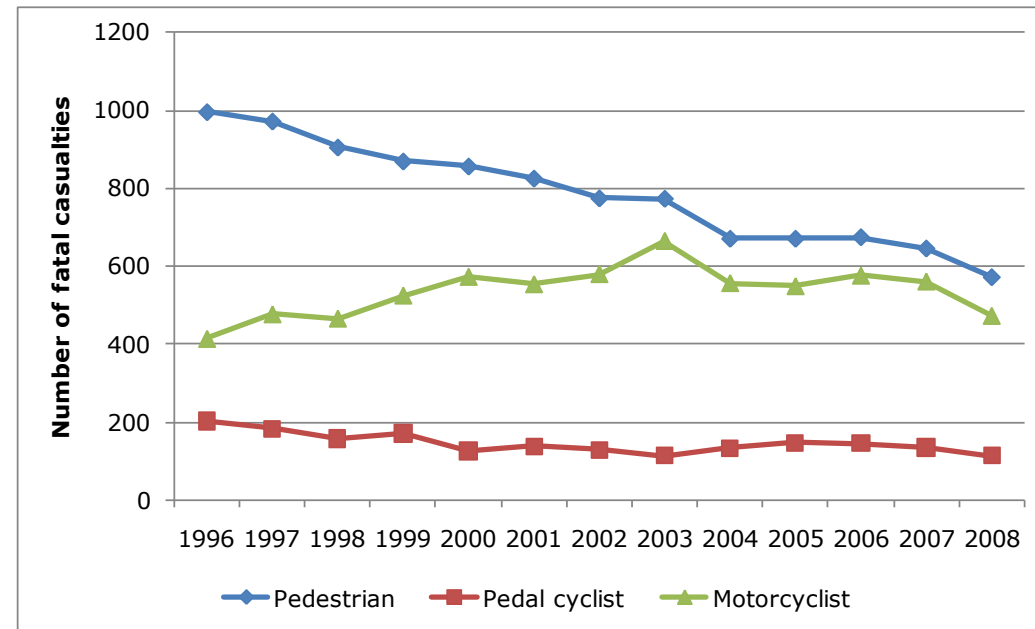
1. The point at which the axes cross should almost always be (0,0)
2. If not, insert “Z” into axis
3. Add group information by using different colours and/or shapes



Continuous data and time series

1. If the line represents point measurements, add points to the lines
2. The optimal maximum number of lines on one chart is 3-4
3. Choose the Y-axis scale carefully

Figure 11: Number of casualties killed in road accidents from 1996 to 2008, by road user type



Key messages when presenting data

- *Think what the message is from the data*
- Clearly label tables etc.
- Use common sense when selecting table/ chart
- Ask yourself whether there is something not quite right

References

- University of Reading Statistical Services Centre. Informative Presentation of Tables, Graphs and Statistics.
www.reading.ac.uk/ssc/publications/guides/toptgs.html accessed 04/08/10
- Albert Goodman. Graphical Data Presentation.
www.deakin.edu.au/~aggodman/sci101/chap12.php accessed 04/08/10
- Local Government Data Unit – Wales (2004). Presenting Data. Performance Management Support Portfolio 5.3.
http://www.dataunitwales.gov.uk/Documents/Publications/adviceandguidance/CPS15010_Presenting_Data_090722_FINAL_eng.pdf accessed 04/08/10
- Department of Chemical and Process Engineering. Presenting Data.
<http://lorien.ncl.ac.uk/ming/dept/tips/writing/data.pdf> accessed 04/08/10.
- Phillip Good and James Hardin (2009). Common Errors in Statistics (and how to avoid them). 3rd edition: Wiley, USA.



**Do You
Have Any
Questions?**