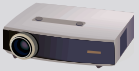


## Module 5

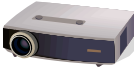
# Communicate information through numbers, graphs and tables



Lecture slide #1-3

**After completing this module, students will be able to:**

- Collect number information to address a problem
- Organise this information in table form
- Analyse the information in different ways
- Present information in different ways
- Critically interpret the information

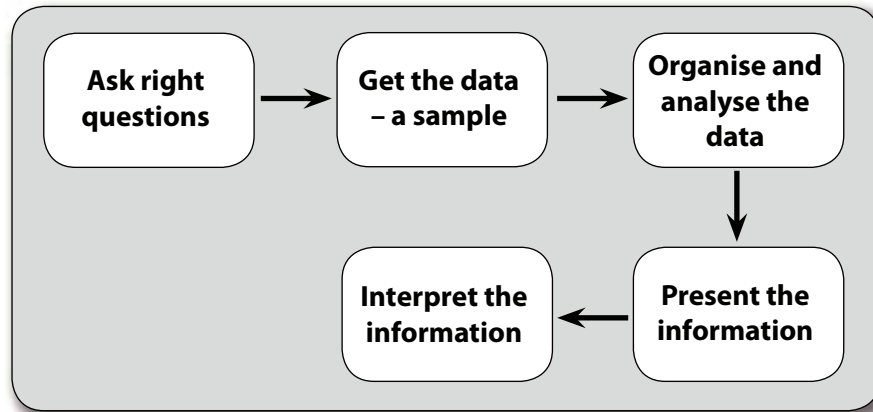


Lecture slide #4-10

## Introduction

Explain to students the process of gathering information, and explain where such information is used e.g. scientific studies or market research.

**Flow of work to find statistical answers to a problem:**



**Method of tackling a statistical problem, i.e. a problem in number form:**

Here is, in more detail, the order in which you have to work.

1. Ask the right *questions*.
2. Decide on appropriate *data* (information) to collect.
3. Establish *availability of relevant data* from libraries or from the government, e.g. see website of the government, which is: [www.statssa.gov.za](http://www.statssa.gov.za); economic data is also available at [www.resbank.co.za](http://www.resbank.co.za)
4. If data is not available, collect own data by careful *sampling* methods
5. *Summarise and analyse* the data, including calculation of average/central values (mean, mode, median) and of spread/range (range, interquartile range, percentiles).
6. *Present the data* using scatter diagrams or linear graphs for paired data, bar graphs, histograms, pie charts or pictograms for discrete data; linear graphs for continuous data.
7. *Look for relationships*, i.e. *interpret* the presented data, be it to find correlation, or differences; *interpretation must be done critically*.

In any working environment, statistics (answers obtained by means of working with numbers) are used to answer questions. In the business environment, statistics are used in production, customer service, marketing, purchasing and supply, human resource management, and finance.

**Mathematical information can be represented in:**

- A simple **flow diagram**:  
Input value  $\rightarrow \times 5 \rightarrow + 3 =$  Output value
- This can also be **written as an algebraic formula**:  
 $y = 5x + 3$
- **Put into words**, this would read:  
A certain number must be multiplied by 5, and 3 must be added to the answer to obtain the output value.

- This sort of **information can lead to a table** when a number of input values are used.

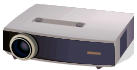
<b>Input values</b>	0	3	6	9	12	15	18	21	24
<b>Output values</b>	3	18	33	48	63	78			

Complete the table - use your calculator.

- Information **can also be plotted as a linear graph** on square-ruled paper or graph paper by using a horizontal number line for the input values and the vertical number line for the output values. The Cartesian plane with coordinate points is used for plotting graphs.
- Graphs (numerical pictures) can take the form of:
  - line graphs,
  - broken line graphs,
  - scatter graphs,
  - bar charts,
  - histograms,
  - pie charts,
  - pictograms,
  - stem-and-leaf diagrams.

*See slides for examples of graphs*

## 1. Collecting the information



Lecture slide  
#11-14

### Introduction

*Explain to students the process of gathering information in questionnaires.*

*The important point to emphasise is that the statistics are used to answer questions. e.g. “Do rural learners generally get lower school marks than their urban counterparts?” or “Do our customers generally prefer chocolate ice-cream to vanilla ice cream?” The process of analysing statistics is to get them to answer a question.*

### 1.1 Accurate terminology and clear questions

*Since we are gathering the information to answer questions, it is vital that in the process of research, we ask the right questions to ensure we get the right answer. Explain to students how questions can be ambiguous e.g. Did you wait too long in the queue doesn't tell us how long the person waited, and so gives us no idea of how long the average person waited in the queue, but it does help to determine whether customers are happy.*

#### Terminology

When investigating unemployment in a country the researcher has to be very clear about what is understood by the word “unemployment”.

“Unemployment” could mean:

- the jobless people in a certain age group who do not have work, or
- only the jobless people who are actively going for job interviews and failing to find work.

*South Africa has two measures of unemployment, the official or narrow definition and the expanded definition. The narrow definition is defined as “people within the economically active population who did not work seven days*

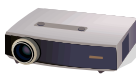
*prior to the interview; want to work and are available to start work within two weeks of the interview and have taken active steps to look for work in the four weeks prior to the interview” The expanded definition drops the criterion of actively seeking work.*

**Questions**

When facing any particular problem, you have to think carefully which questions to ask.

The questions must be:

- relevant, i.e. have a direct bearing on the problem at hand.
- formulated clearly, i.e. must leave no space for confusion.



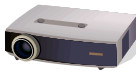
Lecture slide #15

**Activity 1 – The right questions**

*Students can complete this activity for homework or in class.*

1. You need to create a budget for your work. Explain step by step how you would gather information for this budget.
2. You work at a clinic. You have to establish the frequency of occurrence of different illnesses in the community. Set up a table for a tally list. Compile five relevant questions.
3. You are a market researcher for a clothing company. Choose men’s or women’s outfits and then formulate ten questions for the man-on-the-street to establish the possible trend in fashion of the next season.

**1.2 Data collecting tools**



Lecture slide #16-17

*Explain to students the ways of collecting :*

- *Collecting data through the company’s normal operations.*
- *Collecting data by doing your own primary research e.g. surveys.*
- *Collecting data through secondary research e.g. databases.*

*Explain to students the difference between discrete and continuous data.*



**Activity 2 – discrete and continuous data**

*Students can complete this activity in class or for homework.*

Tick in the table which of the data sets are discrete and which are continuous. Read pg. 137 and 138 to complete the last column.

Event	Discrete data set	Continuous data set	Measure of centre – mean or mode
Number of defective items in a production batch	√		Mode
Measurement of time		√	Mean
Number of telephone calls received in a given hour	√		Mode
Measurement of babies’ birth weights		√	Mean
Categories such as “type of tourist”	√		Mode
Numbers of years grouped together	√		Mode
Clothing numbers worn by a number of adults	√		Mode
Measurement of length		√	Mean
Money measurement		√	Mean



Lecture slide #19-21

Explain to students that a survey is a way of collecting data which can include the following methods:

- questionnaires,
- tick/tally lists,
- general questions with individual answers.

Explain what a sample is and why we use samples (samples are covered in more detail later, so it is not necessary to go into too much detail here)

### A. Questionnaires

One major way of acquiring data is by **questionnaires**, i.e. a set of questions relevant to a specific research topic. In any questionnaire the questions should be:

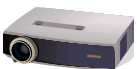
- short – people are more inclined to participate in a survey if it does not take much of their time,
- straight-forward/unambiguous, and
- simple - the answers must be easy to interpret. We do not want a long sentence for an answer, therefore use yes/no (**dichotomous**) answers, **multiple choice** answers, or **numerical rating scales** normally on a scale of 1 to 5 points (e.g. “never” to “always”, or “poor” to “excellent”). This scale is called a **Likert scale** of measurement.



#### Example:

Of a questionnaire in a bank for clients to complete after having been attended to by a sales consultant:

SERVICES RATING	SALES CONSULTANTS	
1. How long did you have to wait?	5 min / 10 min /15 min / longer	
2. Was the consultant friendly?	Yes	No
3. Were all your questions answered?	Yes	No
4. Was the product/procedure explained sufficiently?	Yes	No
5. Was the time taken to handle your transaction reasonable?	Yes	No
6. Will you use our services again?	Yes	No
Any other comments: _____		



Lecture slide #22-23



### Activity 3 – Accurate terminology and clear questions

Students can complete this activity in class or for homework

1. Noluthando did market research before starting her small business and before investing her inheritance into the *Little Laundry*.
  - a. Decide whether she asked sensible questions. If not, re-do the questionnaire for her by giving her some **constructive/positive criticism**.
  - b. Add three more questions to her list.

Market survey of the <i>Little Laundry</i> - Question			Yes	No
Do you have time to do your own washing?				
What time do you get home on week days?				
How much do you spend per month on your laundry?				
Do you like doing your own washing?				
Would you make use of the services of <i>Little Laundry</i> ?				
Would you like your laundry to be delivered?				
How many kilograms of clothes do you wash per month?	10 kg	100 kg	1 000 g	>1 000 kg

2. Compile your own market research questionnaire for Thumeka's *Little Shop*, i.e. try to find out what her clients prefer to buy.



Lecture slide #24

**B. Tick or tally lists are also used in surveys**

*Explain to students how to use a tick or tally list and the advantages of using them. The Post Office example in the text book provides an excellent illustration of this.*

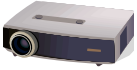
The tally is normally kept as follows: four straight vertical lines for the first four occurrences and the fifth occurrence is a horizontal line crossing the previous four lines. It is then easier to count the eventual totals by merely counting in fives.

You will use such a tick/tally list in the activity which involves a survey on cell phone usage of your friends/co-workers/fellow-students.

Next time you visit the Post Office look out for the pad that is lying in front of the person who serves you. The pad is a weekly planner with and on it provision is made for each hour of the day. The attendant will make a tick in the appropriate hour box every time he or she sees a client. A significant amount of information is gained in this manner. The Post Office uses this from time to time to management to determine the following:

1. Which hour(s) of the day are the busiest?
2. How many minutes it takes on average to serve a client.
3. How hard the attendant is working, the total number of clients served per hour, day and week.
4. Whether they should open more booths for certain times of the day to prevent long queues.
5. When the best time is for an individual to go on lunch.

	08 – 09	09-10	10-11	11-12	12-13	13-14	14-15	15-16	16-17
<b>Mon</b>	+++	+++ +++	+++ +++	+++ +++ +	+++ +++ +	+++ +++ +	+++ +++	+++	+++
<b>Tues</b>	+++ +++	+++ +++	+++ +++ +	+++ +++ +	+++ +++ +	+++ +++ +	+++ +++ +	+++ +++ +	+++ +++
<b>Wed</b>	+++	+++ +++ +	+++ +++	+++ +++ +	+++ +++ +	+++ +++ +	+++ +++ +	+++ +++ +	+++ +++
<b>Thurs</b>	+++	+++ +++	+++ +++ +	+++ +++ +	+++ +++ +	+++ +++ +	+++ +++ +	+++ +++ +	+++ +++ +
<b>Fri</b>	+++ +++	+++ +++	+++ +++ +	+++ +++ +	+++ +++ +	+++ +++ +	+++ +++ +	+++ +++	+++ +++ +
<b>Sat</b>	+++	+++ +++	+++	+++ +++	+++ +++ +				



Lecture slide #24

**Assignment 1 – The gathering of information**

Divide the class into groups of two per group. Each group must decide on a specific SME business venture in their line of interest.

**Create a virtual business by deciding on the following:**

- Are you the owners of the SME?
- How many people are employed in the business?
- What is the company’s annual turnover?
- Which industry sector is the business in?

Your business subscribes to a weekly newspaper, *The Business Gazette*. Once a year this paper includes a survey list on the back page to judge reader sentiment. They entice readers to fill this in by stating that a donation of R10 will be given to a certain charity per received survey list and by putting a few computer program up as lucky draw prizes.

A. Fill in the survey for *The Business Gazette*.

**Reader Survey**

1. How often do you read *The Business Gazette*?

	Every issue	Some issues	This is the first issue	Never
Printed version				
On the web				

2. If your answer in question 1 was “Never”, why have you not read it?

---



---

3. Complete the questionnaire by ticking in the correct spaces:

Subject	Read	Very useful	Sometimes useful	Not useful
Business opportunities				
Success stories				
How to start a business				
Family-business				
Advice on BEE issues				
Franchising				
Taxation				
National economic issues				
Staff and HR issues				
Distribution				
Production				
IT				

Marketing advice				
Export issues				
Case studies of people who started own business				

4. Is there anything that you would want included in future issues?

---

5. Is there anything that you would like changed in *The Business Gazette*?

---

6. Are you:

- Manager of an SME    Owner of an SME    Employee of an SME

7. How many people, including yourself, are employed in the company?

1 – 10	11 – 30	31 – 50	51 – 100	>101

8. What is your company's annual turnover?

< R100 000	R100 000 - R250 000	R250 000 – R500 000	R500 000 – R1 million	>R1 million

9. Personal details

Gender :  M    F

Age :  Under 25  
 25 – 34  
 35 – 49  
 50 – 59  
 60 +

In which industry sector do you work?

(Financial/IT/ Import-Export / Manufacturing / Health / Sales / Construction / Education / Hospitality / Tourism / Entertainment / Catering / Advertising / PR / Retail / Other – please specify)

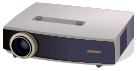
---

B. Organise the results of the whole class into one table by using the answers of only questions 1; 3; 6; 7; 8; and 9.

C. Write a paragraph reporting the conclusions of the survey.



## 1.3 Samples



Lecture slide  
#28-29

*Explain to students why sampling is used.*

*Explain what a representative sample means e.g. includes both rural and urban people.*

*Explain what a random sample is i.e. there is no bias about who is selected.*

## 2. Organising the collected information

After the **collection** of data, it has to be **organised to get a total picture** and better understanding of the information. After the collection and organisation of the data, we can **apply mathematical techniques to make sense** out of, i.e. analyse the data.

The first step in organisation of data is normally the drawing up of a table. Tabulated data is much easier to grasp than written paragraphs.



Lecture slide  
#28-29

### Activity 4 – Organisation of data – role play

*This activity can be completed in class or for homework*

You work as a marketer for a cell phone company and have to do market research. You want to find out in which province of South Africa the population density, i.e. people per square km, is the highest. This will directly determine where you will do most of your marketing.

Look at the following information regarding the regional demographics in South Africa:

- **Eastern Cape**  
- 168 966 square kilometres; 6.9-million
- **Free State**  
- 129 825 square kilometres; 2.9-million
- **Gauteng**  
- 16 548 square kilometres; 9.5-million
- **KwaZulu-Natal**  
- 94 361 square kilometres; 9.9-million
- **Limpopo**  
- 125 755 square kilometres; 5.6-million
- **Mpumalanga**  
- 76 495 square kilometres; 3.5-million
- **Northern Cape**  
- 372 889 square kilometres; 1-million
- **North West**  
- 106 512 square kilometres; 3.4-million
- **Western Cape**  
- 129 462 square kilometres; 4.7-million



#### Did you know?

##### EASTERN CAPE

**Capital:** Bisho

**Major city:** Port Elizabeth

**Languages:** 83.4% isiXhosa, 9.3% Afrikaans, 3.6% English

**Population:** 6 919 071 (2006)

**Share of total SA population:** 14.6%

**Area:** 168 966 square kilometres

**Share of total SA area:** 13.9%

**Population density:** 41 people per square kilometre

**Gross regional product:** R88-billion (2003)

**Share of total SA GDP:** 8.1%

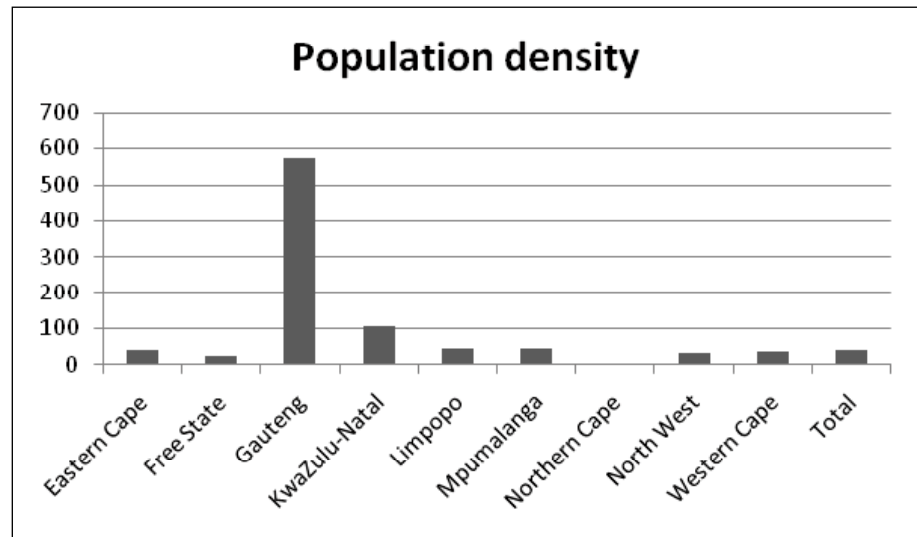
#### Questions

You might want to work through the whole module before doing questions 3 – 6.

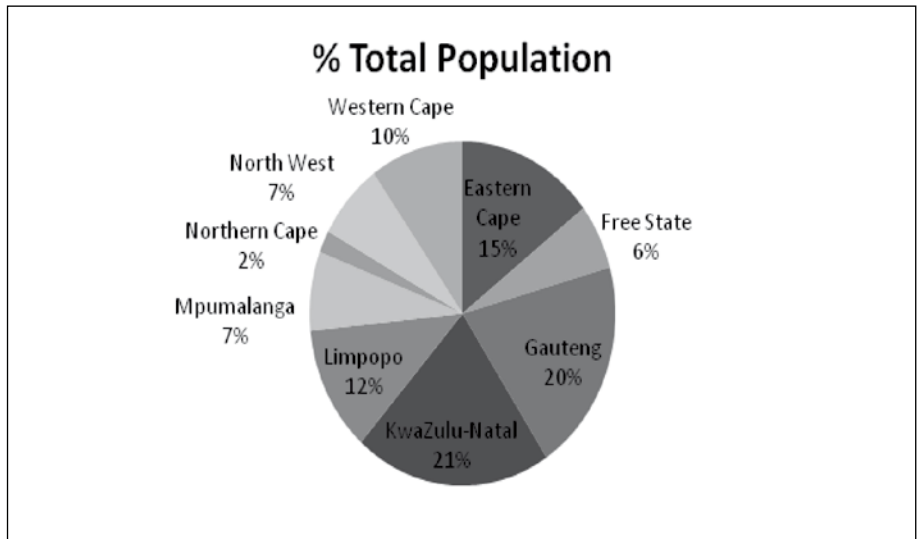
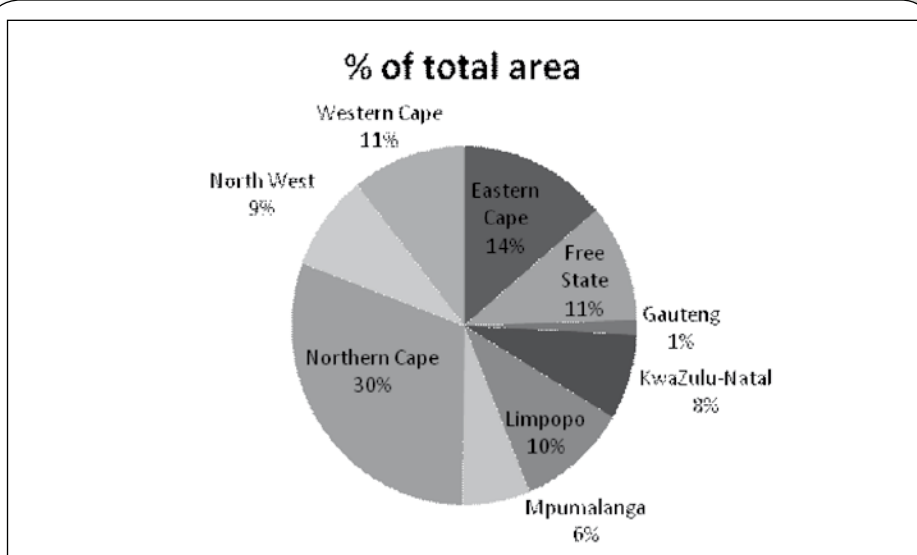
1. For the Eastern Cape the % of total area has been calculated as well as the % of total population and the population density per square km. Now do the same calculations for the other regions.
2. Tabulate all of the data.

Province	Area (km <sup>2</sup> )	% Total Area	Population (million)	% Total Population	Population density
Eastern Cape	168966	13.84%	6.9	14.56%	41
Free State	129825	10.63%	2.9	6.12%	22
Gauteng	16548	1.36%	9.5	20.04%	574
KwaZulu-Natal	94361	7.73%	9.9	20.89%	105
Limpopo	125755	10.30%	5.6	11.81%	45
Mpumalanga	76495	6.27%	3.5	7.38%	46
Northern Cape	372889	30.54%	1	2.11%	3
North West	106512	8.72%	3.4	7.17%	32
Western Cape	129462	10.60%	4.7	9.92%	36
Total	1220813	100.00%	47.4	100.00%	39

3. Draw a bar graph with the regions set on the horizontal axis and the population density on the vertical axis.



4. Is it possible to represent all of the data on one bar graph?  
*No, you are dealing with different sets of axes*
5. Could a pie chart be used instead of the bar graph?  
*For some of the graphs, not for population density*
6. Select one column of information that you consider could be presented in a pie chart. Draw such a pie chart.



7. In which province is the population density the highest?

*Gauteng*

8. Which is easier to understand:

- the list of raw figures given you,
- the table of organised and analysed data,
- the bar chart,
- or the pie chart?

Rate these different ways you used to present the information on a scale of 1 – 5. Create a table for this rating exercise. Give reasons for your decisions.

**Case study 1 – Thumeka’s Little Shop**

Remember Thumeka’s *Little Shop*?

During her first four months she kept a tally of her sales as presented in the table below.

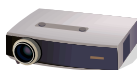
Item	Week 1	Week 2	Week 3	Week 4	Month 2	Month 1
Toffees	46	59	40	46	191	185
Lollipops	36	42	37	42	157	156
Boiled sweets	50	46	59	65	220	215
Potato crisp packets	12	15	10	15	52	52
Cheese curls	32	32	30	34	128	124
Chocolate bars	9	13	12	9	43	37
Tennis biscuits	10	17	22	17	66	58
Tins of carbonated drinks	18	16	21	19	74	55
Fruit juice 250 ml containers	25	28	37	35	125	120
Bananas	28	30	32	36	126	120/month
Oranges	24	30	34	42	130	340/month
Match boxes	28	32	20	19	99	15 boxes / day
Cigarette type 1	5 packs of 20	6 packs	4 packs	5 packs	20 packs	25 cigs. / day
Cigarette type 2	4 packs of 20	5 packs	6 packs	6 packs	21 packs	18 cigs. / day

**Questions – Thumeka’s Little Shop is open for five days each week.**

1. The previous table indicates the number of items sold during the weeks of the second month for Thumeka’s Little Shop. Insert the correct number of tally lines (in brackets) for the chocolate bars and Tennis biscuits.
2. Organise the information for the first two months in a separate table by referring to the module on finance (summative assessment).
3. Compare the totals of the two months and decide whether Thumeka’s business is looking good.
4. Use any method to present any of the data and draw a conclusion from your presentation.



Lecture slide #39-43



Lecture slide  
#44-46

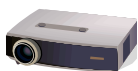
### 3. Analysing the organised data

#### 3.1 Percentages

*The use of percentages should be revision to most students as this was covered earlier in the course as well as in mathematical literacy 2.*

Percentages are used in almost every statistical report. It is a way of expressing a proportion and of making comparisons.

1. Converting a percentage into a fraction:  
“Per cent” means per hundred, i.e. a fraction with the denominator being a hundred. E.g.  $5\% = 5/100$  and  $35\% = 35/100$ .
2. Similarly, a fraction can be turned into a percentage by multiplying by 100.



Lecture slide  
#47-48

#### Activity 5 – Group assessment

*Students should complete this activity for homework.*

1. A student at a college smokes 17 cigarettes per day on average.
  - a. Find out how much one packet of 20 cigarettes cost and then calculate how much he spends per month on cigarettes.
  - b. If he earns R4 500 per month after deductions, calculate what percentage of his salary he spends on cigarettes.  
*He is spending 12,72% of his R4500 earned on cigarettes.*
2. If Thumeka makes R3 946 during a certain month at her Little Shop, and she spends 45% of her money on food, calculate how much money she spends on food?  
*R1775,70*
3. Create your own two percentage sums, one of each kind. Work with a friend on this one. Present this answer to the class and the class has to critically assess the calculations.

#### 3.2 Proportions and rates

*The use of proportions and rates should be revision to most students as this was covered earlier in the course as well as in mathematical literacy 2.*



Lecture slide  
#49-51

#### Activity 6 – Group assessment of class’ proportions and rates calculations

*Students should complete this activity in groups in class or for homework.*

1. Thumeka wants to decide which of her salty products is the best seller for the two months she has been open.
  - a. Calculate the proportion between “Crispy chips” and “Cheese curl” sales.  
*Crispy chips : Cheese curls = 124 : 52 in month 1 and 128 : 52 in month*
  - b. Calculate what percentage of her salty food sales are made up of “Cheese curls”.  
*The percentage of cheese curls with respect to “salty” sales =  $52 \div 176 \times 100 = 70\%$*

2. Noluthando spent R2 650 on food during a certain month when she earned R5 620.
  - a. Calculate the proportion of her money that went into food.  
 $2650 : 5620$
  - b. Calculate the percentage of her money that went into food.  
47% of her money spent on food
  
4. Create your own proportion calculation. Work with a friend on this one. Present this answer to the class and the class has to critically assess the calculations.



Lecture slide #52-53

### Activity 7 – Interpret the news – Wages

Read the following excerpt from the Sunday Time’s “Business Times”, ([www.btimes.co.za](http://www.btimes.co.za)) of 18 April 2004.

*Students can complete this activity in class. It is a good activity to use for a formative assessment.*

#### **Wage increases set to beat inflation. “Salaries” by Marcia Klein.**

Inflation and wage settlements have been running neck and neck this year, but wage increases are set to outstrip inflation by the end of the year.

According to Andrew Levy & Associates’ quarterly wage settlement survey, inflation and the average level of settlement were both at 9,7% in the first six months, but it is expected that the settlement level will overtake inflation before the end of the year.

Settlements have ranged from 6,8% to 15% and the average minimum wage was R1 650 in the first six months. Initial union demands ranged from 8% to 146%, while management counter offers ranged from 0% to 11%. In just over half the cases, wage negotiations were settled without industrial action, but in 16,1% of cases there was strike action.

- a. Would a settlement close to 10% be fair in light of the above paragraph?  
*Yes, it would keep up with inflation rate – the average wage settlement is given as 9,7%*
- b. If the settlement level overtakes the inflation level, which will be the highest?  
*The settlement level.*
- c. If a worker earns R1 750 per month and the wage settlement is 10%, what adaptation can the worker expect?  
*He could expect R1925 which is an adaptation of R175.*
- d. Will it be possible for a worker to earn R1 000 per month after these wage adaptations have been implemented?  
*Not probable as the minimum wage is given as R1650.*
- e. Use the last paragraph of the excerpt to explain how wage negotiations normally take place. Negotiations take place and the workers bring their expectation down and the employer raises their % offer until an agreement is reached.



Lecture slide  
#54-56

### 3.3 Measures of centre

Measures of centre include the **mean, mode** and **median**.

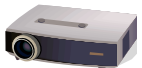
*Explain why we use statistics i.e. to give as a clearer picture of what is happening, and why we will use more than one measure of statistics to get a clearer idea of what the data is telling us.*

*Explain how to calculate the mean, median and mode. If you have a scientific calculator, it may be worthwhile to get it to generate random numbers to allow for examples.*

*Explain how an outlier or a stray may affect the mean.*

#### Examples of averages

- Class average for a test.
- Average number of days that students were absent from college.
- Average temperature in December of Windhoek.



Lecture slide #57

#### Activity 8

*Students can complete this activity in class. It shouldn't take them more than five minutes.*

*State whether you would use the mode or the mean for the information supplied in the table in Activity 2.*

*See activity 2*

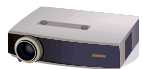


Lecture slide  
#58-62

### 3.4 Measures of spread

*Explain to students the following concepts:*

- **Range:** *The difference between the largest and smallest observations.*
- **Interquartile range:** *The difference between the upper quartile and the lower quartile*



Lecture slide  
#63-66

#### Activity 9 - Work in pairs

- a. Calculate the **mean, mode, median, range and interquartile range** of the following two sets of data which represent numbers of cigarettes smoked per day:

A for women :13; 10; 19; 8; 17; 5; 12; 17; 7; 13; 15; 12; 9; 13; 20

$$\text{Mean} = 190 \div 15 = 12$$

$$\text{Mode} = 13$$

$$\text{Median} = 13$$

$$\text{Range} = 20 - 5 = 15$$

$$\text{Interquartile range: } 17 - 9 = 8$$

B for men: 18; 8; 16; 14; 10; 18; 17; 17; 20; 11; 17; 22; 12; 21; 13; 21; 15

$$\text{Mean} = 270 \div 17 = 15,88$$

$$\text{Mode} = 17$$

$$\text{Median} = 17$$

$$\text{Range} = 22 - 8 = 14$$

$$\text{Interquartile range: } 20 - 13 = 7$$

- b. Compare the results in a paragraph.

*For the A group of data - The mean, mode and median are all close to 13, and the 15 values range from 5 to 20. The interquartile range*

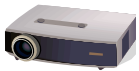
*shows most values to lie between 9 and 17. For the B group of data – The mean, mode and median are all close to 17, and the values range from 8 to 22. The interquartile range shows most values to lie between 13 and 20. The B group seems consistently a bit higher than the A group.*

2. The wages in a fruit processing factory are as follows (numbers in brackets represent the number of people in the specific wage group):

2 350 (2)	2 500 (5)	2 750 (4)	3 000 (2)	4 500 (1)
12 740 (1)	15 800 (1)	2 600 (2)	3 250 (3)	2 000 (2)
5 000 (1)	3 500 (2)	1 800 (2)	2 800 (2)	

**Calculate for the data:**

1. The mean  
 $107390 \div 30 = R3579,67$
2. The mode  
*R2500 (4 occurrences)*
3. The median of wages in this factory  
*The median of wages in this factory = R2750*
4. The range of the wages  
*The range of the wages = 15800 – 1800 = R14000*
5. The interquartile range  
*Quartile 1 = 7,75 position = R2500*  
*Quartile 2 = 23rd position = R3250*  
*Interquartile range = 3250 - 2500 = 750*



Lecture slide #67-68

## 4. Presentation of results

The organised results can be presented to the reader in many ways.

Presentations are useful to:

- summarise data and facilitate conclusions,
- present information concisely to an audience, both of which will,
- save time during discussions.

In the introduction to this module students read about:

- Bar chart, also called bar graph
- Histogram
- Straight line graph
- Broken line graph
- Line of best fit on a scattergraph
- Pie chart

The three types of presentation used most frequently are **line graphs (also broken line graphs), bar graphs and pie charts.**

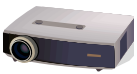


The type of graph you choose will depend on whether you want to:

- Summarise or describe information - stem-and-leaf plot, bar graph.
- Compare information - back-to-back stem-and-leaf plot, compound bar chart.
- Describe a relationship between two variables – straight line graph.

## 4.1 Line graph

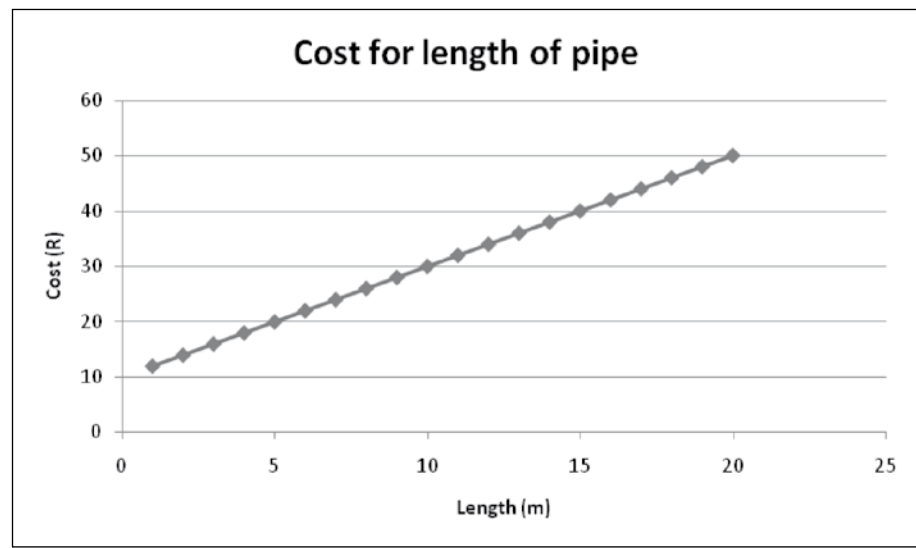
Students covered drawing a line graph in module 3. This section should for revision for them.



Lecture slide #69-70

### Activity 10 – Line graphs – self assessment

A hardware shop charges for irrigation piping with the formula:  
 Cost of the pipe system = 2 times length in metres of pipe plus R10 cutting fee. Sketch a line graph of the information. Refer to the module on patterns and relationships.



## 4.2 Scattergraph with line of best fit

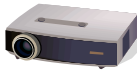
Scatterplots are the most used graph in research, where the relationship between two variables is investigated. Many of the statistical techniques for scatterplots are not covered here, however, students should be able to identify a trend and draw an approximate trendline.

### Example

Economists plot the relationship between inflation rate and the unemployment rate. On graph paper the one variable is then plotted from the horizontal axis (X-axis) upwards, and the other variable is plotted from the vertical axis (Y-axis) towards to the right. These scatter plots can then provide us with information about the relationship between the two variables.

Students should be able to identify the following:

- Linear
- Non-linear
- No relationship



Lecture slide  
#71-76

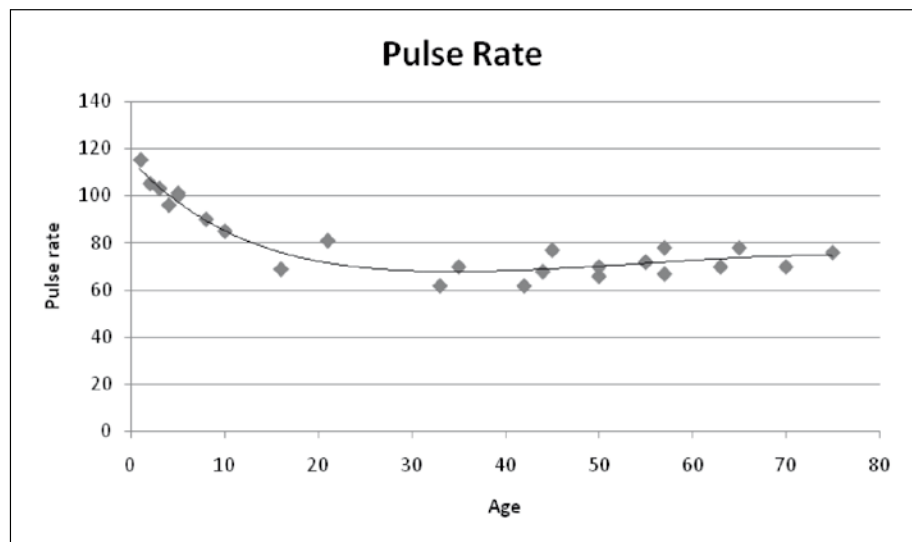
**Activity 11 – Draw a scatter diagram**

Compile a scatter diagram of the following information (Graham – Statistics):  
The **age (top row)**, and **pulse rate in beats per minute (bottom row)** of a sample of 25 people are given.

10	3	4	5	8	1	16	21	5	2	33	35
85	103	96	101	90	115	69	81	100	105	62	70

Age (top) and pulse rate (bottom row) – continued:

55	50	45	65	57	70	42	75	57	44	63	50	55
72	66	77	78	78	70	62	76	67	68	70	70	72



**Questions**

- How would you describe the relationships for the category:
  - from 0 to 15 years old, and
  - from 30 to 70 years old?

*For the younger age group the trend is downward i.e. the older the child the slower the heartbeat. From 30 to 70 years old the trend is upward, i.e. the older you are the faster the heart rate.*
- Now add two lines of best fit for the two categories in the scatter gram.
- What might you deduce from the lines of best fit?  
Heart beat is fast just after birth, but becomes slower up to about 20 years old. Thereafter the heartbeat increases slowly but surely year by year
- Add your own pulse rate (count pulse for 10 seconds and times by 6) to the scattergram.
- Why could it be necessary to include the phrase “pulse rate on arising after a night’s sleep” in a comparison of age to pulse rate?  
*Pulse rate varies throughout the day, is slowest and more comparable just before rising in the morning.*
- Which other factors could affect pulse rate?  
*Level of stress; Smoker / non-smoker; Level of fitness; Illness*



Lecture slide #77-79

### 4.3 Bar chart

Bar chart	Histogram
Frequency measured by height of bar.	Frequency measured by area of the bars.
Used for discrete data sets.	Nearly always used for continuous data sets.

The main features of the **bar graph/chart** are the following:

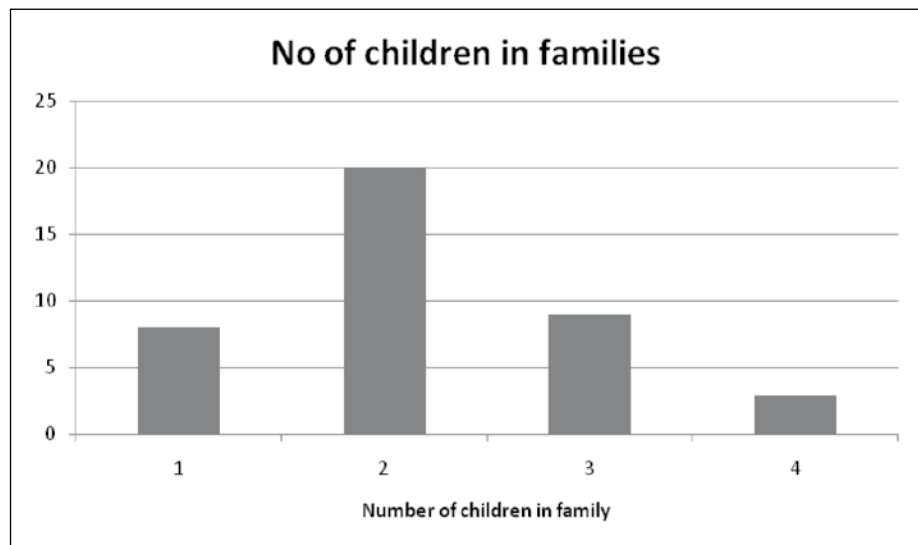
- The *width of the bars* must be exactly the same – this allows a fair comparison between the data sections as the eye of the reader only sees the height of the bars to establish the difference between the sections.
- The *distance between the bars* must be exactly the same. The gaps show that the *horizontal axis is not continuous* but represents separate sections.
- The vertical and horizontal *axes must be labelled*.
- The graph has to be informatively *titled*.
- A bar chart can also lie horizontally and each bar can represent more than the one aspect of the information under investigation (*compound bar chart*).
- A *component bar chart* (each bar is split to reveal its components) should ideally have no more than 5 or 6 of sections / components to still read easily. This is similar to a *pie chart* where the sections of the pie represent components of the total.



#### Activity 12 – Bar graph

Students should complete this activity in class or for homework in groups of 2.

1. Draw a bar graph of the information in the table.

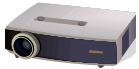


2. What is the modal class of the data?  
The modal class is two-child families.



Lecture slide #80-81

Number of children per family	Number of families in category
1	8
2	20
3	9
4	3



Lecture slide #82

## 4.4 Pie chart

Statistical data can be represented as slices or sectors of a pie or circle, the angles of the slices being in proportion to the frequency of occurrence or the percentages of the data sectors.

- The **information presented in the pie chart must add up to a whole**, e.g. the sections of the whole day spent on different activities – adding up to 24.
- The angles of the circle are calculated according to a simple formula:

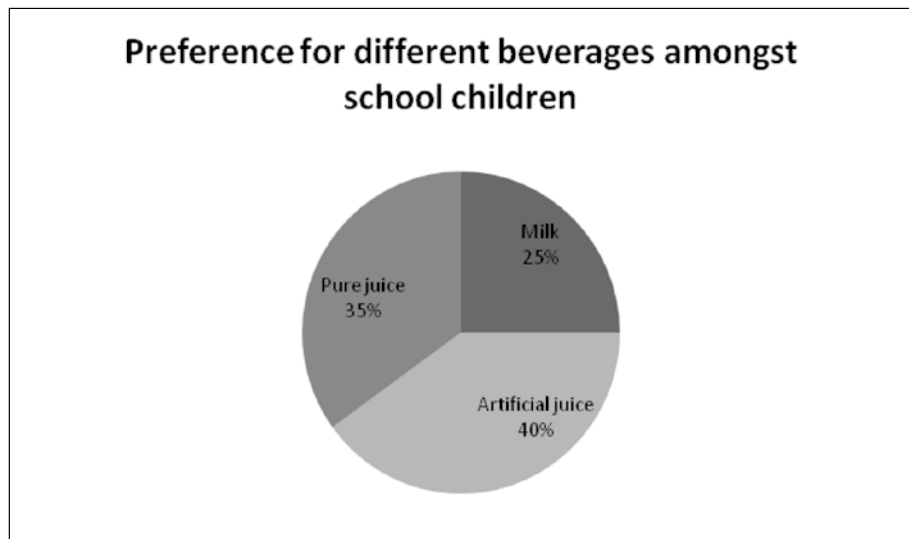
$$\text{Angle of sector} = \frac{\text{Frequency}}{\text{Total number}} \times 360^\circ \quad (\text{one full circle} = 360^\circ).$$



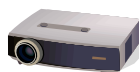
Lecture slide #83-84

### Activity 13 – Pie chart

1. Sketch a pie chart of the information in the last column.



2. Put percentages on the pie chart.
3. Add a title to the sketch.
4. Write a sentence as conclusion to the sketch.



Lecture slide #85

**Activity 14 – Choose the right representation for the event**

*Students can complete this activity in class. It shouldn't take more than five minutes.*

The two variables	
months of the year, against ... rainfall figures on vertical axis	<i>Bar chart</i>
distance travelled by truck/bus, against ..... cost of transport/travelling	<i>Straight line graph</i>
shoe sizes, against ..... number of persons on vertical axis	<i>Bar chart</i>
different companies, against ... expenditure on vertical axis	<i>Bar chart</i>
Countries, against ... % of adults with HIV on vertical axis	<i>Bar chart</i>
share price index, against ... time	<i>Broken line graph</i>
area of house/road, against ... building cost	<i>Straight line graph</i>

Numerical information is gathered in order to answer questions and make judgements.

You have now seen that the order of work is:

- collect data,
- draw up a table (i.e. organise and summarise data into a table),
- draw a graph (i.e. summarise or represent data in a graph).



Lecture slide #86-99



**Activity 15**

*Students can complete this activity in class. They will need at least two lessons for this activity. It can be used as a formative assessment tool.*

**In the news – Cell phone statistics**

Most adults and many children in South Africa have their own cell phones. Many of these people just use their phones to contact friends. A cell phone can also have a safety purpose. But, for many of us, cell phones are a necessity for work purposes. The following comes from an article by Frik Els in the magazine, *FinWeek* of 6<sup>th</sup> March 2008.

“The 55 000 delegates attending the recent World Mobile Congress in Barcelona, Spain, had reason to be optimistic. More than **1,1 bn mobile phones were sold last year – 160 m more than in 2006** - the fastest growth in five years, according to GFK research house. **Every sixth person on the planet bought a mobile phone in 2007**. This year the outlook is no different, the subprime mortgage and credit crises, the global slowdown and energy crunch notwithstanding.”

However: “In the fourth quarter of 2007 one US company’s mobile devices business swung to **a loss of US\$1,2 bn from a profit of \$2,7 bn**. This company, Motorola, is now the number three producer of cell phones worldwide, with **13% of the market share**. That is down from owning 26% of the global market the year before. Its stock is at a four-and-a-half year low and many analysts are calling for the mobile division to be spun off to turn the company around. The prime reason for Motorola’s woes is said to be that it cannot compete in emerging markets with the likes of Nokia (the Finnish company sold more than 400 m phones in 2007) and Samsung, which overtook Motorola last year as world number two with 15%.

China, India and Africa are where the growth lies for handset manufacturers. Compared with close to 100% in developed markets (Italy is **the most saturated market, with 112% cell phone penetration**), the emerging world is far from saturated: Mobile phone penetration in China is less than 40%, **in Africa it is 27%**, while India lags with **fewer than one in every five people** owning a mobile phone. That is despite sales doubling on the sub-continent to 90 m in 2007.

For Samsung, growth in emerging markets has been even more spectacular. The South Korean company saw an **increase in sales last year in the Middle East and Africa of 80%** and on average **emerging markets increased by half** versus a still healthy **21% in developed markets**. This year the company hopes to sell more than **200 m units against 161 m the year before.**”

“... SA follows both developed and developing market trends. For example, because candybar or block phones are most associated with entry-level products, it is no surprise that in sub-Saharan Africa as a whole they account for 85% of all phones sold, while in SA it is just over half. Similarly, 96% of phones sold in sub-Saharan Africa are thicker than 13 mm, while slim, fashion phones account for 18% of the SA market. More than 50% of SA cellphones have a built-in camera. In Uganda it is less than 3% and in Europe 85%.”

### Questions

1. What is the percentage increase in sale of mobile phones from 2006 to 2007?

*The percentage increase in sale of mobile phones from 2006 to 2007 is 1100million – 160 million increase amount = 940 million sales amount for 2006. Percentage increase = 17%*

2. What is the probability that a person has a cell phone, globally speaking?

*The probability that a person has a cell phone, globally speaking, is one out of six or 16,67%.*

3. For the Motorola company, which is the negative figure, the US\$1,2 bn or the \$2,7 bn?

*For the Motorola company, the \$1,2 billion is the negative amount as it represents a loss.*

4. What is the percentage decrease in business swing for Motorola?

*The percentage decrease in business swing for Motorola, is 21,47 negative.*

5. What is the percentage drop in the global market share of Motorola?

*Motorola has 13% of the global market share. This down from owning 26 % therefore the drop was 50% (not 13% - this would be the answer to how many percentage points of the market share did Motorola lose)*

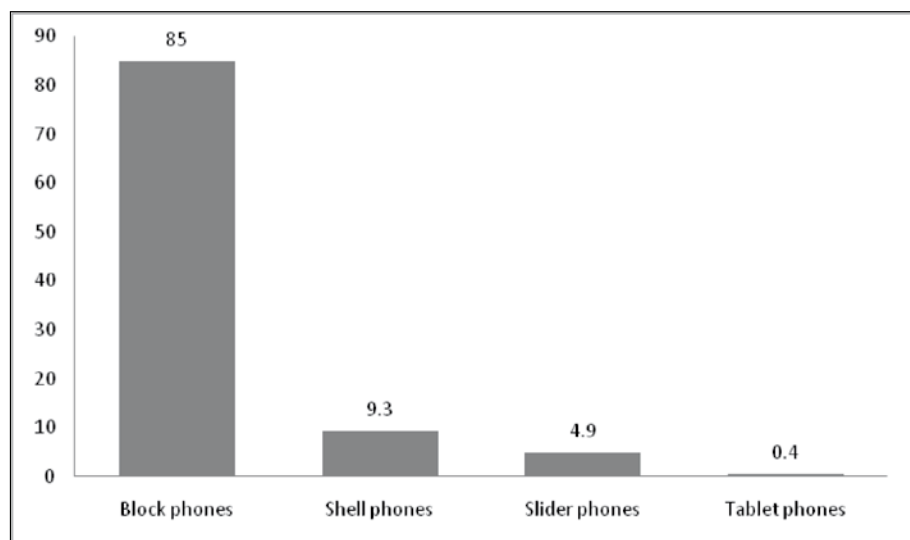
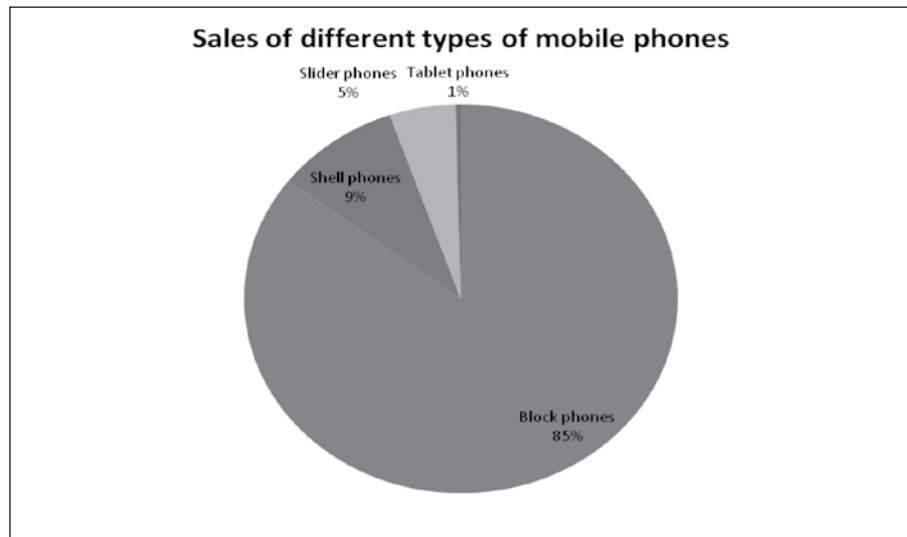
6. Explain what “112% saturation in cell phone penetration” means.

*“112% saturation in cell phone penetration” means that for every 100 people there are 112 cell phones.*

7. What was Samsung’s sale figure on the sub-continent (India) in 2006?

*Samsung’s sale figure on the sub-continent (India) in 2006, has increased by half of 80% which is an increase of 40% on the previous year’s sale*

8. What percentage increase does Samsung expect in 2008?  
*Samsung hopes to sell 200m minus 161m units which is 39m units.  
 The percentage increase in 2008 will then be 39 divided by 161 times 100 = 24,22%*
9. In sub-Saharan Africa the following phones are mainly sold: block phones (85%), shell phones (9,3%), slider phones (4,9%) and tablet phones (0,4%).  
 Draw a pie chart as well as a composite bar graph of this information.



10. Look at the sketches of bar charts for cell phones sales in Africa and for Europe. Why are the scales of the vertical axes the same on both graphs and why do they both reach up to 300 million?  
*The scales of the vertical axes are the same on both graphs and both reach up to 300 million to enable comparisons.*
11. From these bar charts, calculate the number of people living in Africa and in Europe.  
*Africa - 27% is 250 million, therefore 100% = 926 million; Europe - 95% is 800 million, therefore 100% = 842 million*

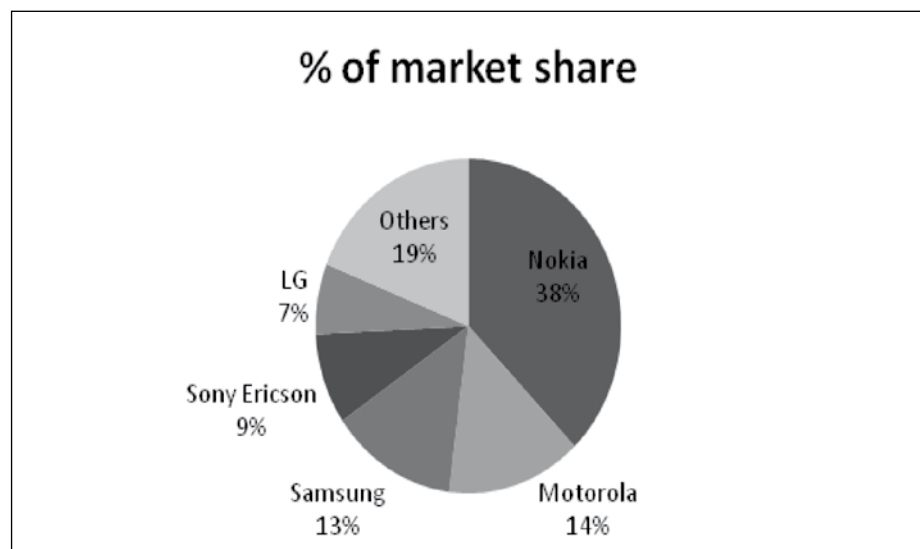
12. “In Asia those users who may have started on a US\$30 to \$40 phone may upgrade in 2008 to something twice the cost.” Convert these US\$ figures to South African Rand, also called ZAR.

*If the conversion factor is : R1 = \$0,1247 then \$30 = 30 ÷ 0,1247 = R240,58 and \$40 = R320,77*

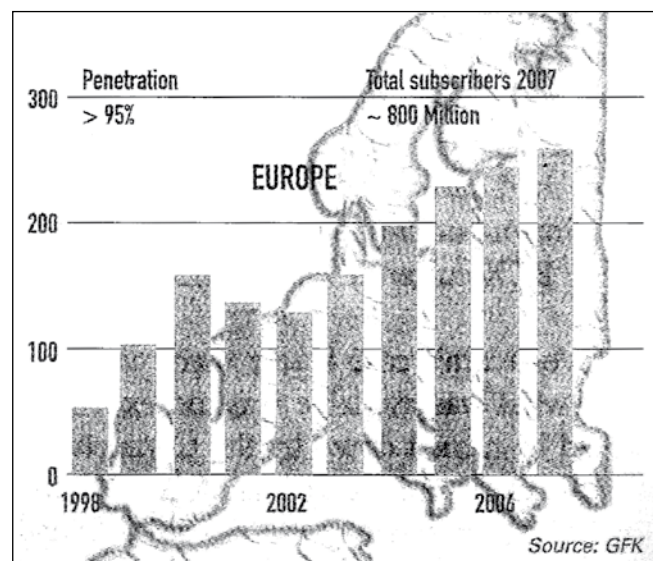
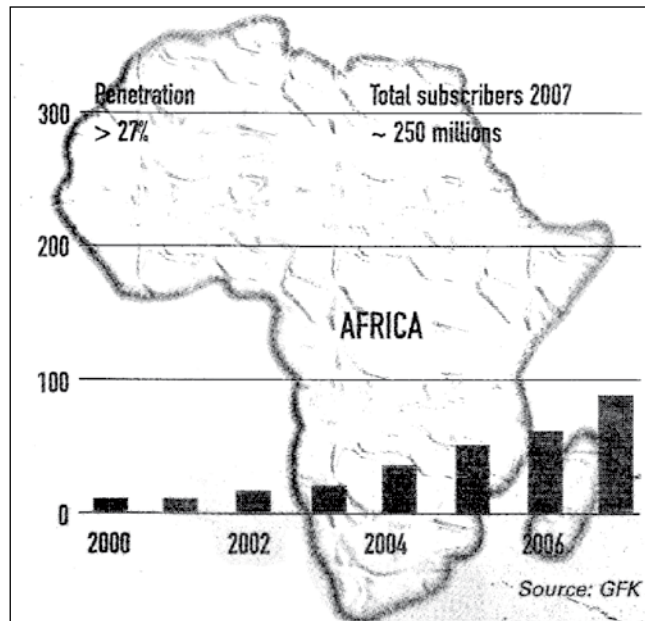
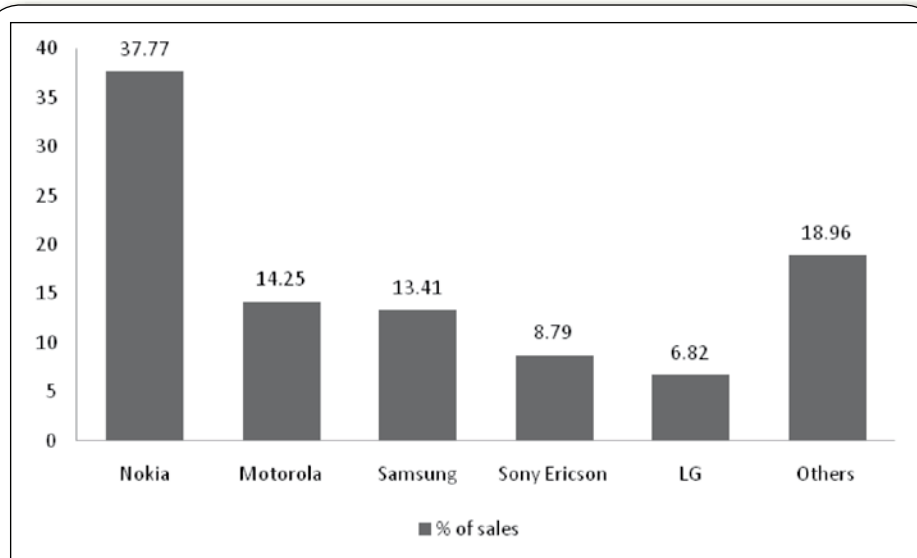
13. For the following table:
- Calculate the % share of sales of each company.
  - Calculate the % variance of market share between 2007 and 2008 and mention whether the variance is positive or negative.
  - Write the total sales for 2007 in words.
  - Draw a bar chart as well as a pie chart of the 2007 market shares and include the percentage of sales on each bar or pie chart sector.

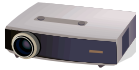
**Worldwide mobile terminal sales in 2006 & 2007 (thousands of units):**

Company	2007 sales	2007 - % of market	2006 sales	2006 - % of market	% variance in market share
Nokia	435 453	37,77%	344 915	34,81	26,25
Motorola	164 307	14,25	209 250	21,12	21,47
Samsung	154 540	13,41	116 480	11,76	32,68
Sony Ericson	101 358	8,79	73 641	7,43	37,64
LG	78 576	6,82	61 986	6,26	26,76
Others	218 604	18,96	184 588	18,63	18,43
Total					







Lecture slide  
#100-101

### Activity 16– Collect information on cell phone usage – two people per group.

*Students should do this research in their own time out of class and hand in a report giving their findings. It is recommended that they work in groups of four and be given at least two weeks to complete this assignment.*

1. Do a survey amongst your friends, family or fellow students/workers to investigate cell phone usage. Try to cover all age groups. Groups must not survey the same people.

#### **You have to find out:**

In which age group does the participant fall?

Which brands of cell phone are used?

Which kind of cell phone is preferred, i.e. slide, swivel or block?

Is there any distinction between male/female cell phone usage?

Were the phones bought cash or on contract?

Do the participants believe that their contract is the best?

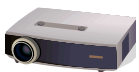
How much money is spent per month on airtime?

Do the participants use the cell phone only for SMS messages, for incoming calls only, or for all calls and SMS messages?

#### **Instructions:**

- a. Design your own survey form to establish the above facts and anything else that you might consider important.
- b. Survey at least 20 people per group. Inform all participants that the survey is anonymous.
- c. Count the tally marks where applicable and then use tables to organise the data.
- d. Analyse the data by calculating percentages, mean values and range where applicable.
- e. Draw a table on the white board and include the information from all of the groups of the class (now more or less 200 surveyed individuals).
- f. Once again analyse the data, now for the whole class.
- g. Compare your analyses to that done for the class and draw conclusions.
- h. Interpret your results in a few sentences.
- i. Draw a bar chart of cell phone brand statistics from the whole class.

## 4.5 Misleading the reader



Lecture slide #102

*Explain to students the way the graphs can mislead readers, for examples excluding certain data.*

*Most importantly students need to understand that correlation does not imply causation (cum hoc ergo propter hoc).*

*An example is given by Wikipedia:*

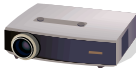
*“In a widely-studied example, numerous epidemiological studies showed that women who were taking combined hormone replacement therapy (HRT) also had a lower-than-average incidence of coronary heart disease (CHD), leading doctors to propose that HRT was protective against CHD. But controlled trials showed that HRT caused a small and significant increase in risk of CHD. Re-analysis of the data showed that women undertaking HRT were more likely to be from higher socio-economic groups (ABC1), with better than average diet and exercise regimes. The two were coincident effects of a common cause, rather than cause and effect as had been supposed.”*

Students must therefore be careful to ask: “Did A cause B; did B cause A or did C cause A and B?” if two events are correlated.

### Mistakes on graphs

To avoid misleading the reader when sketching graphs, take note:

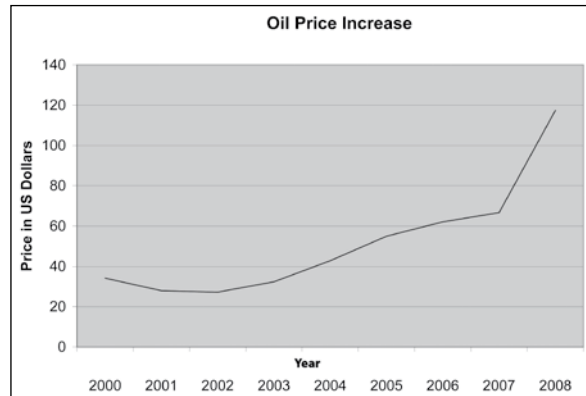
- Graphs are there to save time, but a mistake in the scale of the x- and y-axes, by accident or on purpose, can totally mislead a reader.
- The axes are two number lines and run from the bottom left-hand corner.
- These lines run from small numbers (usually zero, if numbers are only positive) to larger numbers – up for the Y-axis and to the right for the X-axis.
- You **have to start the scale at zero** on the axes, or otherwise indicate clearly that the particular scale has been interrupted.
- Remember that the small blocks on the axis must represent equal values of whatever is being represented (uniform scale). You **have to stick to a uniform scale on both axes**.



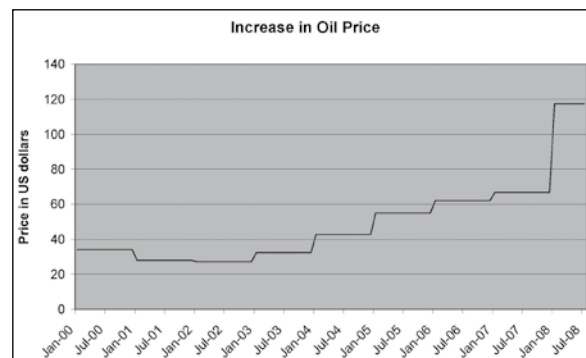
Lecture slide  
#103-104

### Activity 17 – Manipulating graphs to suite the situation – class discussion.

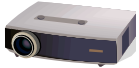
1. Compare and discuss the similarities and differences between the two graphs.



2. Identify the mistakes in the representation of data.



Both graphs are correctly plotted and represent the same data with similarly named axes. The only different is the way the data has been plotted. The first graph is smoothed between the points, while the second graph seems to join the points directly.



Lecture slide #106

**Other errors (intentional or unintentional) resulting in the misuse of statistics**

- *Confusing questions* can be asked (See example in Introduction).
- *Sampling can be incorrect or inadequate*, e.g. the absentee rates of a firm could be estimated by sampling the absentee rates only every five working days; or a household’s budget can be estimated by sampling bills and payments of only the first week of every month.
- *Analysis of data can be misleading*, e.g. percentages can be confused. Look at the following example from *Teach yourself Statistics* by Alan Graham.

Status of household	Number of children in situation of poverty	% children
Parent in full-time work	1025	34
Lone parent	938	31
Unemployed parent	763	25
Pensioner	320	10

The message is that the largest number of children in poverty is from households where there is a parent in full-time work. This is not wrong but does not tell the full story, and does not reflect what we are interested in - only households in poverty have been used in the calculation.

However, if another column is added to the table, indicating the total number of children in the population that are from the different kinds of households, it now becomes a meaningful calculation. The percentage of each separate group of households in poverty are compared with the total number of households, no matter its economic state. It becomes clear that only 11% of the children from full-time worker households are in poverty (calculated 1 025 divided by 9 330 times 100).

Status of household	Number of children in situation of poverty	Total number of children in population	Percentage of children in situation of poverty
Parent in full-time work	1 025	9 330	11
Lone parent	938	1 250	75
Unemployed parent	763	930	82
Pensioner	320	470	68



Lecture slide #106

**Activity 18 – Percentages - self assessment**

Complete the last column in the table above.



Lecture slide #106

## 5. Probability mathematics

Probability in maths is the attempt to deduce/predict the outcome or result of an event.

- If a perfectly balanced coin is flipped, it is logical to expect that the outcome heads and the outcome tails are equally likely. Thus the probability of heads appearing is  $\frac{1}{2}$  or 50%.
- In general then, if an experiment has  $n$  possible outcomes that are all equally likely, the probability of any particular outcome to occur is  $1/n$ .

In statistical calculations, probabilities are expressed as:

- Numbers between zero and one, expressed either as a fraction or as a decimal.
- A percentage, between 0 and 100 per cent.

Thus an unlikely event might have a probability of 0,1 whereas a 50% chance/ even chance would have a probability of 0,5 and a definite event would have a probability of one or a 100%.

### A. Theoretical probability

When outcomes are equally likely, it is fairly straightforward to calculate their probabilities.

You simply divide the number of ways in which that outcome can occur by the total number of possibilities.

Thus, a **formula to calculate probability** where outcomes are equally likely is:

$$p = n/N,$$

where  $p$  is the theoretical probability of a particular outcome occurring,  $n$  is the number of ways that outcome can occur and  $N$  is the total number of possible outcomes.



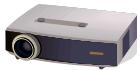
#### Activity 19

*Students need to work in pairs. This activity can be completed in class and shouldn't take more than a couple of minutes.*

Event	Possible outcomes	Probability of outcome occurring
A coin tossed	1. heads 2. tails	0,5 or or 50%
Playing card chosen	1. a spade 2. a heart 3. a diamond 4. a club	(13/52) which is one quarter = 25% or 0.25
Result of an examination	1. passing 2. failing	0,5
Gender of a baby	1. a girl 2. a boy	0,5
Swimming across the Atlantic Ocean in a day	Impossible	0
That you will die some day	Definite	1

#### Questions

1. Fill in the rest of the last column by using the formula for theoretical probability.
2. For which of these events is the given probability possibly not true?
2. Probability possibly not true for "results of an examination" as it depends how much the student has studied.



Lecture slide #113

## B. Other probability situations

More practical situations do not lend themselves to this *simple theoretical approach*.

1. The **relative frequency approach** to probability expresses an outcome's probability as its long-run relative frequency of occurrence.
  - For example, if 600 out of the last 1 000 customers entering a store have made purchases, the probability that any given customer entering the store will make a purchase is approximately  $600/1000 = 0,6$  or 60%.
  - The larger the sample size, the more accurate will be the estimate of the desired probability.
2. The **subjective approach to probability** is largely based on intuition.
  - **Example:** A producer wants to launch a Broadway musical and wants to predict the probability that the show will return a profit to investors. He has to keep in mind the reputation of the specific musical, the quality of other current Broadway shows, the state of the economy, and the fact that only about 25% of all Broadway musicals are profitable.

## C. Experimental determination of probability



Lecture slide #114-115

### Activity 20 – In groups of two

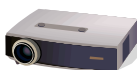
*In order for this activity to work, the class needs to standardise on one experiment. Due to availability, coins would be easiest. Students should initially work in pairs, but the later stages should be done as a class activity.*

Work in pairs, one person tosses or chooses, and the other one keeps the tally.

- toss a coin 100 times (note heads or tails), or
- choose a playing card 100 times (note hearts, clubs, diamonds or spades, or
- toss a die 100 times (note 1,2,3,4,5, or 6).

This is **an experimental simulation**.

1. Calculate the ratio/proportion of e.g. heads counted to total number of throws.
2. See if the probabilities calculated from these experiments are the same as that which you would expect from the given formula,  $p = n/N$ .
3. Now add the tally for all the groups in the class, and work out the experimentally determined probability once again.
4. See if the probabilities are now closer to or further away from the expected probabilities calculated with the formula,  $p = n/N$ .  
(The expected theoretical probability of any of the numbers appearing should equal  $1/6$  or  $0,16667$ .)
5. When the number thrown is taken on the horizontal axis of a **bar graph** and the probability of  $0,16667$  on the vertical axis, a sketch of a **probability model** appears. Sketch this model - no gaps between the bars for this sketch.



Lecture slide #114-115

### “Odds” is another way of expressing probability

- In a horse race the odds on each horse are a measure of how likely each horse is to win a particular race.



Lecture slide  
#117-121

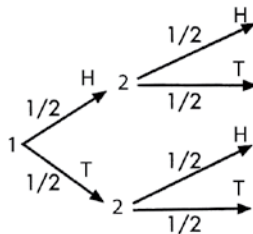
### Activity 21 – Odds and probability

Students can complete this activity in class or for homework.

1. The odds on three particular race horses are as follows:

Horse name	Odds
Mogollon	5 : 1
Isibomvu	20 : 1
Vonk	2 : 1

- Which horse is considered by the bookies as least likely to win?  
*Isibomvu which has long odds*
  - Which horse is considered by the bookies as most likely to win?  
*Vonk which has short odds.*
  - If a punter places R50 on Isibomvu, and the horse wins, how much money would he take home?  
*He will take home  $50 \times 20$  plus the original stake of 50 = R1050.*
  - If he places the same bet on Mogollon, how much money will he take home?  
*R300*
  - If the odds on Vonk had been 3 : 2, and the horse wins, what would a punter win who had placed R20 on the horse to win?  
*Vonk had been 3 : 2 is the same as 1,5 : 1 and he will win  $20 \times 1,5$  plus 20 = R50*
2. Look at the following tree diagram which is a model of the probabilities of obtaining heads or tails when tossing a coin and interpret the probabilities given as fractions.



*Probability to throw two heads in a row will be  $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$*

*This same calculation follows for each route i.e. for heads followed by tails, for Tails followed by heads and for two tails in a row.*

3. A lucky dip consists of 36 envelopes, only 4 of which contain a prize. On the first dip, what is the probability of winning a prize? Suppose the first envelope chosen did contain a prize, and the envelope is removed from the bag. What is the probability of the second dip also winning a prize?  
 $\frac{4}{36}; \frac{3}{35}$



### Summative assessment

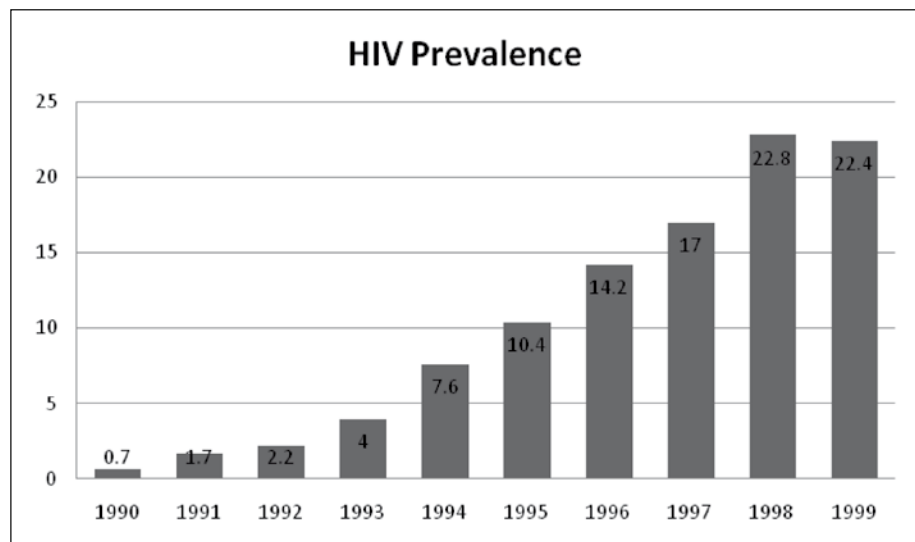
*This forms the final activity before the summative assessment test. Students should complete this activity in class in silence. It can be used as a formative assessment tool, and would be particularly useful to give to students before a test to see advise them which areas to work on.*

- The following table presents results of an HIV survey by the department of Public Health Services amongst women attending anti-natal clinics (i.e. 15 – 49 year-old pregnant women):

HIV prevalence amongst child-bearing women in South Africa  
1990 – 1999.

Year	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999
HIV prevalence (i.e. % of HIV positive women attending anti-natal clinics )	0,7	1,7	2,2	4,0	7,6	10,4	14,2	17,0	22,8	22,4

- Draw a bar chart of the data with the years on the horizontal axis and the % prevalence on the vertical axis. (See the main features of a bar chart – above).



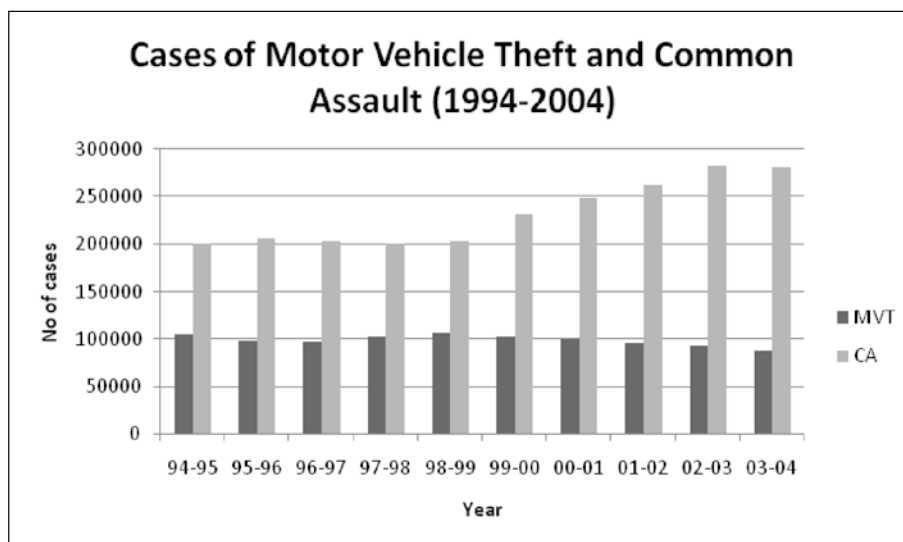
- Make a deduction from this bar chart.
- Now see whether your deduction reads something like this:  
It is easy from the bar chart to make the deduction that: “Estimates of the national HIV prevalence for the years 1990 to 1999 show a steady increase from 0,7% to 22,8% in 1998. This indicates a percentage increase of more than 30 times from the beginning of the epidemic.”
- Have you got any information about the sample size used in the example?  
*No.*
- Would it be possible for an interested person to obtain information about the sample size?  
*Yes such a person could go to Public Health Services.*
- How was the information grouped?  
*Yearly or annual groups.*
- How was the information collected?  
*Information was collected at ante-natal clinics.*



h. How neutral/random was the information collection process?  
 Only pregnant women of child-bearing age were used in the survey. In this group it was only the group who went to clinics who were used. The sample is therefore not random but has to be used as very little information is available.

2. The table presents data on motor vehicle theft as well as common assault cases in RSA from 1994 to 2004.

- a. Write the information on motor vehicle theft in a paragraph.  
*In students own words. It will show that that it is much longer to interpret than a glance at a table or graph.*
- b. Draw bar charts of the information.



- c. Which method, i.e. words, table or bar chart give the best picture of the information.  
*Probably the graph.*
- d. Describe the trends in your own words, i.e. make your own conclusions about the two sets of data.

**Total cases recorded in RSA for theft of motor vehicles and for common assault** ([www.issafrica.org](http://www.issafrica.org))

Year (R)	94 - 95	95 - 96	96 - 97	97 - 98	98 - 99	99 - 00	00 - 01	01 - 02	02 - 03	03 - 04
Number of cases of motor vehicle theft	105 867	98 669	97 332	102 571	107 448	103 041	100 030	96 869	93 133	88 144
Number of cases of common assault	200 248	206 006	203 023	201 317	203 678	232 024	248 862	261 886	282 526	280 942

3. *AIDS affects 5,2 million South Africans.*

(Business Day, November 29, 2005).

“The (actuarial) society’s latest figures come from its new AIDS demographic model ASSA 2003.

It shows there were around 530 000 new HIV infections between mid-2004 and mid-2005 and about 340 000 AIDS deaths during the same period.

The society’s model predicts that 5,8 million South Africans will be infected with HIV by 2010.”

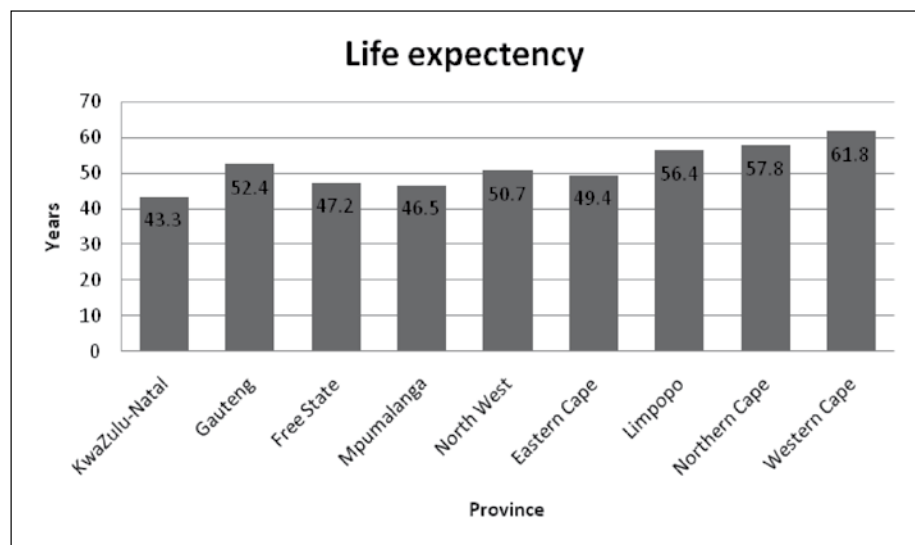
The table of provincial indicators in 2005 comes from [www.tac.org.za](http://www.tac.org.za).

Province	Total HIV (thousands)	Total HIV (prevalence)	15 - 49 HIV (prevalence)	Life expectancy
KwaZulu Natal	1 520	16%	26%	43,3
Gauteng	1 370	14%	22%	52,4
Free State	380	14%	22%	47,2
Mpumalanga	440	13%	22%	46,5
North West	470	12%	20%	50,7
Eastern Cape	630	9%	17%	49,4
Limpopo	380	7%	12%	56,4
Northern Cape	60	7%	11%	57,8
Western Cape	250	5%	8%	61,8
South Africa	5 200	11%	18%	51,0

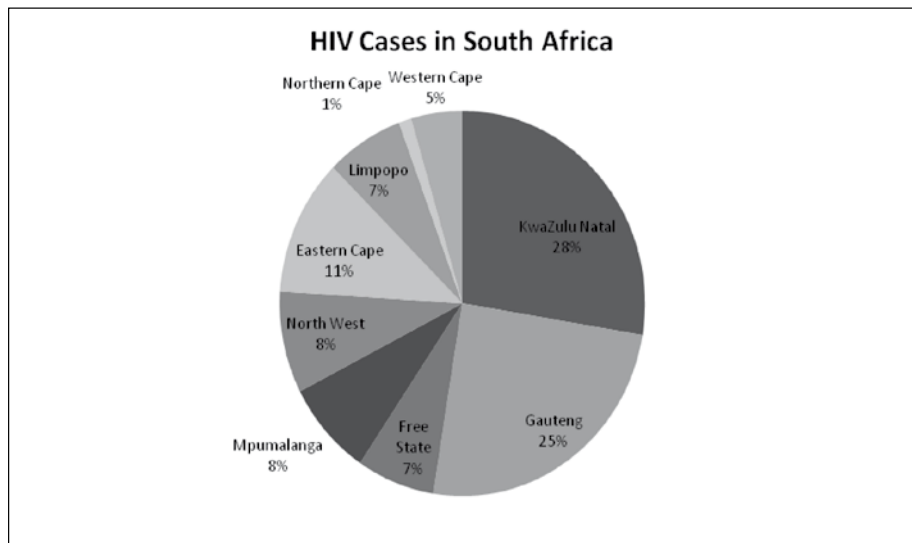
- a. Re-calculate the figures for the bottom row. Do your answers differ from the given data?

*The total for the first column should read 5 500 and the mean for the last column 52 years average.*

- b. Draw a bar chart of data in the last column.

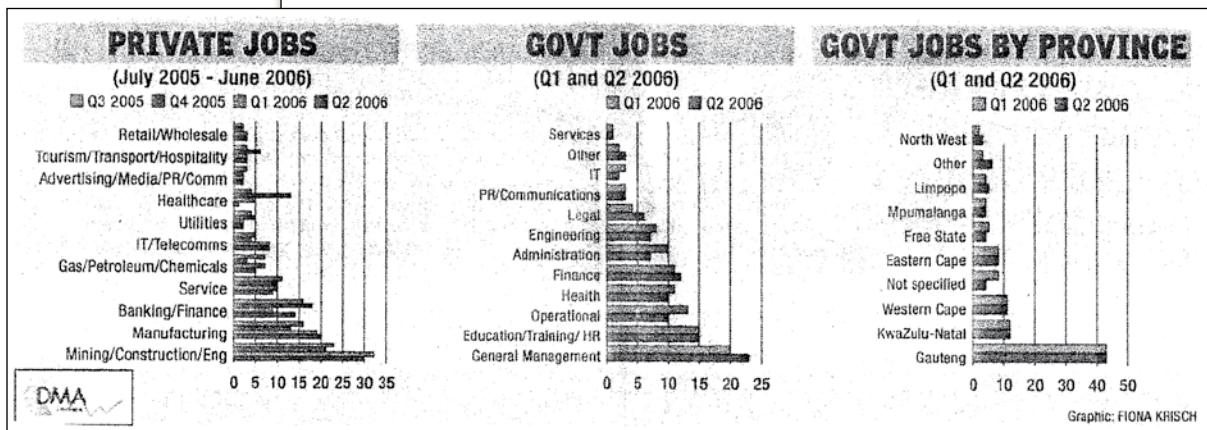


c. Draw a pie chart of data in the first column.



- d. What do you think is meant by “Total HIV prevalence”?  
*It means the percentage of people in that province with HIV.*
- e. Was the information collected here fact or opinion?  
*It was fact.*
- f. What is the probability that a 15 – 49 year old adult in Mpumalanga will get the disease? 22%

4. The bar graphs represent job availability in South Africa. Answer the questions:



- a. Try to figure out what the variable on the horizontal axis is?  
*Percentages.*
- b. What do you think Q1 and Q2 mean?  
*Quarters of the year*
- c. Which two categories of work have the largest job availability in the private sector bar graph?  
*Mining and Manufacturing.*
- d. Which two categories of work have the largest job availability in the government sector?  
*Management and Education.*
- e. Explain your answers in the previous two questions.  
*Evident from the two bar charts.*

- f. In which province are you the least likely to find a job?  
*North West.*
- g. How could data for this survey be gathered?  
*By surveying job ads.*