STAT 6108 :
Official Statistics and Structural Equation Modelling
Second Term, 2019-20

PART 1- OFFICIAL STATISTICS
PowerPoint - 5

(Slide #7, Rectified)

Prof Frederick W H HO

January /February2020

Some points to note in describing situations with *statistical* data (some of which are, in fact, applicable to other kinds of data)

>> 1) A comparison where the base is very small may not be meaningful and can be very misleading (e.g., on the first day of operation of a business, the no. of customers is 5 and one day later, the no. is 7.

Avoid making the statement " … has increased tremendously, by as many as 40% ."

>> 2) Periods under comparison :

Comparison of a current period figure with a previous period figure --
So often, people only say "MMM has increased by 5.4 % (say)",
without mentioning what the basis of comparison is,
i.e.whether this quarter is compared with the last one,
or this quarter is compared with the same quarter last year

(Also, note the subtle differenece between the use of the term
"Rate of Increase " vs " Rate of Growth " [or "*growth rate* "])

Some points to note in describing situations with statistical data (Cont'd)

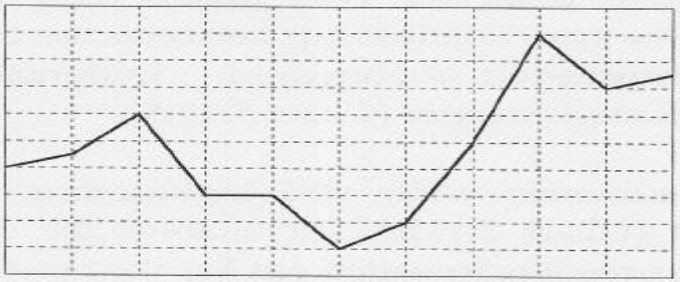
>> 3) X=100 at time t(0) and X=500 at time t(1).
It has increased ***4*** times from t(0) to t(1) (not ***5 times***)

A quantity can increase by (say) 2 times [200%] over a period of time (BUT it cannot decrease by two times !! In decreasing by 100 % the quantity has already become zero. Equally, one quantity cannot be two times *smaller* than another.)

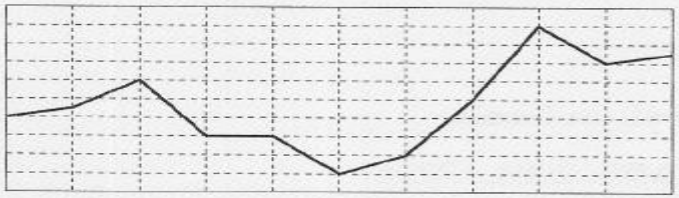
>> 4) From 100 to 120, a quantity has increased by 20 %.
When it drops back from 120 to 100, it decreases by 16.7%
([(100/120) - 1] x 100 %), NOT 20% .

Chart 5.12 **Changing the Visual Image : Contracting or Expanding Vertical (Amount) Scale or Horizontal (Time) Scale Tends to Change the Visual Picture**

Original Scale Arrangement

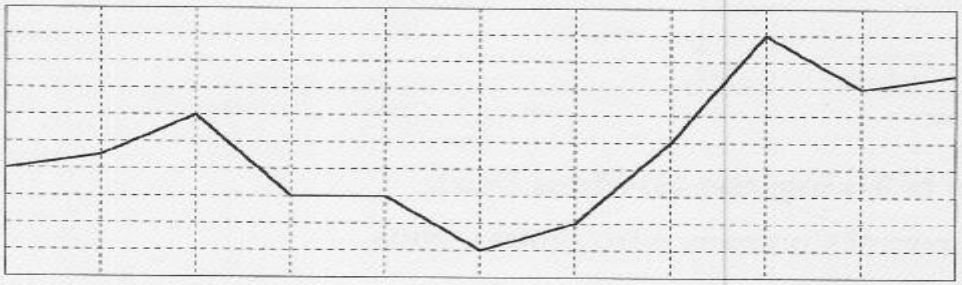


Contracting Vertical Scale



Expanding Vertical Scale

Expanding Horizontal Scale



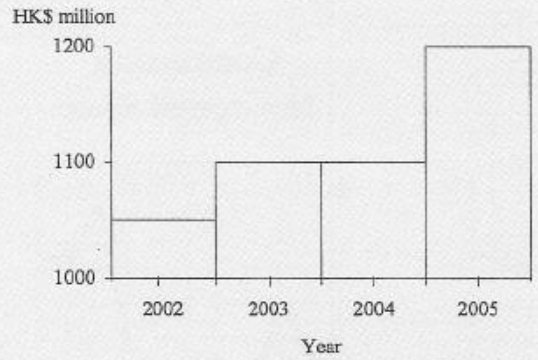
Contracting Horizontal Scale

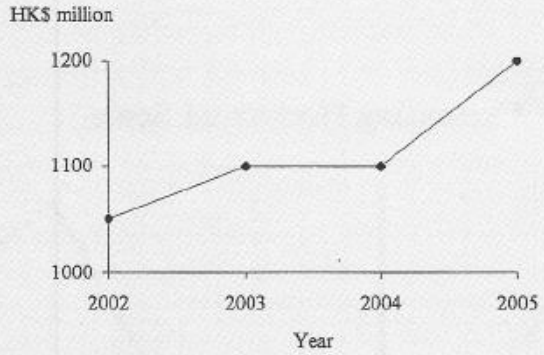
Proper use of the Scale

Chart 5.11

(A) Charts Not Using the Broken Scale

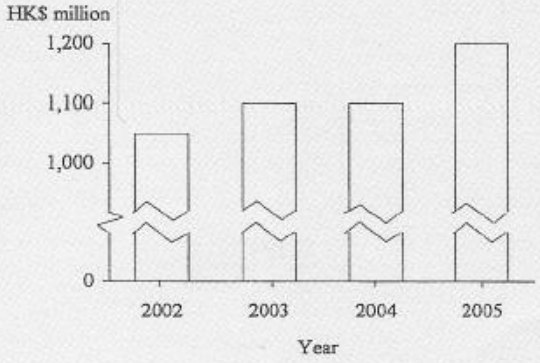
Sales Turnover of Company ABC, 2002 – 2005

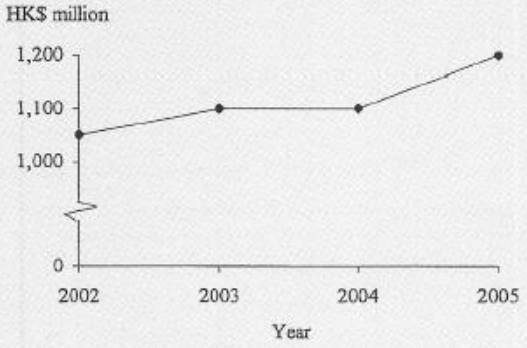




(B) Charts Using the Broken Scale

Sales Turnover of Company ABC, 2002 – 2005





Proper use of the Weighted Average

Example:

Population[1], 2001 Population Census

Area	Male	Female	Sex ratio (No. of males per 1 000 females)
Hong Kong Island	631 854	703 615	898
Kowloon	998 754	1 025 225	974
New Territories	1 650 658	1 692 388	975
Total	3 281 266	3 421 228	959

Note : [1] Excluding the marine population.

It is incorrect to derive the sex ratio for Hong Kong by adding the individual ratios for different areas and dividing by 3. This would give a wrong value of 949 [(898 + 974 + 975) / 3]. The true value of 959 may be calculated only from the total numbers of males and females [(3 281 266 / 3 421 228) × 1 000].

Per cent and Percentage Points

(Slide #7, Rectified)

- The sales value in January was HKD 1.0 Mn . In February it was HKD 1.1 Mn. There was an increase of 10 per cent.

- The sales value in February was HKD 1.1Mn . In March it was HKD 1.24 Mn. There was an increase of 13 per cent.

- The monthly rate of increase was higher in March than in February, by 3 **percentage points** [[13-10]]

……

- The proportion of elderly persons in the population was
  8.7 per cent in 1991 and 11.2 per cent in 2001.
  There was an increase of 2.4 **percentage points**. [[11.2 -8.7]]

De-seasonalisation

(See also a note on an illustration of compiling the seasonal index)

	Original data	Seasonal Index	Deseasonalised data
1st Quarter, Year 1	X(1,1)	S(1)	Y(1,1)
2nd Quarter, Year 1	X(1,2)	S(2)	Y(1,2)
3rd Quarter, Year 1	X(1,3)	S(3)	Y(1,3)
4th Quarter, Year 1	X(1,4)	S(4)	Y(1,4)
1st Quarter, Year 2	X(2,1)	S(1)	Y(2,1)
2nd Quarter, Year 2	X(2,2)	S(2)	Y(2,2)
3rd Quarter, Year 2	X(2,3)	S(3)	Y(2,3)
4th Quarter, Year 2	X(2,4)	S(4)	Y(2,4)
1st Quarter, Year 3	X(3,1)	S(1)	Y(3,1)
(etc.)	(etc.)		(etc.)

Seasonal Index for each of the quarters can be obtained from a standard statistical method

$Y(i, j) = X(i, j) / S(j)$

Reading a table of descriptive statistics

--- to note the use of "rates"

	Unhappy		*Happy*		*Total*	
Age group (Young)	130	26.5%	*100*	*25.0%*	*230*	*(25.8%)*
Age group (Middle)	260	53.0%	*230*	*57.5%*	*490*	*(55.0%)*
Age group (Old)	100	20.4%	*70*	*17.5%*	*170*	*(19.1%)*
	490	100 %	*400*	*100 %*	*890*	*(100 %)*

Fictitious data. Figures may not add up due to rounding

Which is the most "unhappy" group?

Many people only look at (or only have figures on) that part of the above table NOT in italics and conclude that the "middle" age group is the most "unhappy" group (53.0%).

"Rate" of being unhappy for persons in different age groups:

- Age group (Young) = 130/230 = 43.5%
- Age group (Middle) = 260/490 = 53.1%
- Age group (Old) = 100/170 = 58.9%

Thus, "Old" is actually the most "unhappy" group.

Further points to note in the use of statistics

- Note the terms (and the processes) of
  Projection, Forecasting and Prediction
  --there is nevertheless no firm concensus on the exact use of these three words.)

>> Projection – what are likely to happen under existing trends, policies and arrangements

>> Forecasting – what are likely to happen taking into account that people may drastically change behaviour and authorities may change policies and arrangements in the light of developments (or in anticipation of problems)

>> Prediction - forecasting by whichever means (scientific forecasting modified by some "informed guesses" of experienced people; fortune-telling; etc.)

Further points to note in the use of statistics (Cont'd)

Study of the "association" (or "correlation") between variables

("CORRELATION" —which may *not* be Causal Relationship [cause and effect relationship]).

Length of ears and length of life

	Long Life	Short life	Total
Long Ears	N(1)	N(2)	T(LE)
Short Ears	M(1)	M(2)	T(SE)
Total	T(LL)	T(SL)	GT

1. We **tend to** observe, and remember, only those N(1) people; and conclude that people with long ears live longer.
2. Actually, we **need to** look at all the figures above in order to determine whether there is a predominance of people in either N(1) or M(2) and then to conclude whether there is correlation between length of ears and length of life.

Further elaboration of certain issues in official statistics(1)

Recalling Unemployment Statistics:

>To look up HKMDS or HKADS or Report on General Household Survey for a *precise definition of the "unemployed person"*

>"Underemployed Person" –is an "*employed person*" who works less than 35 hours a week and is seeking more work.

Further elaboration of certain issues in official statistics(2)

Recalling GDP:

Growth rate of the GDP

a) One can look at the growth rate of the GDP at current prices (growth over the same period last year) – also termed the "*nominal* growth rate".

b) If the growth rate is on the basis of constant prices, it is also termed the "*real* growth rate "

c)In some economies (not many, but notably USA), the "real growth rate" for a calendar quarter is worked out by comparing the de-seasonalized figures for two adjacent quarters and annualizing it on a compound basis.

Further elaboration of certain issues in official statistics (3)

- HES and CPI

Collection of data in the HES—

The 12-month period (the base period) is split into sub-periods of 2 weeks each

The diary method–

Each household is to keep a detailed diary on the expenditures made on a daily basis within the 2-week period assigned to the household which has been selected into the sample

Further elaboration of certain issues in official statistics (4)

CPI and Inflation

(A) The Rate of Change of the CPI of a certain month from the CPI over a year ago is generally referred to as the INFLATION RATE

(B) To *study the price trend*, one may

i) look at the inflation rate of adjacent months and remark whether there is any acceleration in price increase (say); or

ii) look at the de-seasonalized series and compare the CPI figures of adjacent months [in HK it is recommended to look at a 3-month moving average (centred at the middle month) in the de-seasonalized series with some adjacent 3-month moving average figure]

(End of PowerPoint – 5)

[[End of Part I of Stat 6108]]