Statistical literacy guide
Index numbers

Last updated: September 2007
Author: Paul Bolton

What are index numbers?

Index numbers are a way of presenting proportionate changes in a common format. Their most common use is looking at changes over time, but they are not restricted to this and can be used to compare differences between local areas, groups of the population etc.

Indices can be constructed from 'simple' data series that can be directly measured, but are most useful for measuring **changes in quantities that can not be directly measured** normally because they are the composites of several measures. For instance, the Retail Prices Index (RPI) is the weighted average of proportionate changes in the prices of a wide range of goods and services. The FTSE 100 index is the weighted average of changes in share prices of the 100 largest companies listed on the London Stock Exchange. Neither of these measures could be expressed directly in a meaningful way.

Index numbers have **no units**. Because they are a measure of change they are not actual values themselves and a single index number on its own is meaningless. With two or more they allow proportionate comparisons, but nothing more.

How are they calculated?

The methods used to construct composite indices such as the RPI are complex and include their coverage, how/when the base year is changed (re-basing), the relative weighting of each constituent measure, frequency of changes to weighting and how these are averaged. In such cases the base value may also reflect the relative weights of its component parts. However, it is helpful to look at the construction of an index from a 'simple' data series as it helps to better understand all index numbers.

Indices that look at changes over time need a **base year**; those that do not, normally use an average figure for their base. If the raw data is a simple series of values then the base year/figure is **normally set to 100** and the rest of the series multiplied by the same number (100 divided by the base value) to get the index numbers. Hence a series 50, 60, 80 becomes 100, 120 (60 x 100/50), 160 (80 * 100/50) if the first figure is set at the base value.

What do they mean?

100 is normally chosen as the base year/figure as it makes for the simplest comparisons. Just take 100 away from the figure you are interested in and you are left with the percentage change. The figures below could be a time series or a non-time series index. Value B is the base value and equals 100.

A	B	C
77	100	123

If this were a time series then it is very clear that, compared to B, the value was 23% less in period A (73-100= -23) and 23% more in period C (123-100= 23). Similarly if this series looked at geographical variations with B as the national average then it is also clear that region A had a value 23% below average and region C 23% above.

The choice of base year is very important as it affects those comparisons that can most easily be made and hence the usefulness of any particular index.

The next example shows how index numbers can be used to compare series. Here 1990 is set at the base year and values are given for three series A, B and C. They might be presented in index form because their actual values are not easily comparable (eg GDP for a range of large and small economies) or their units are different (eg carbon dioxide emissions, energy use and GDP).

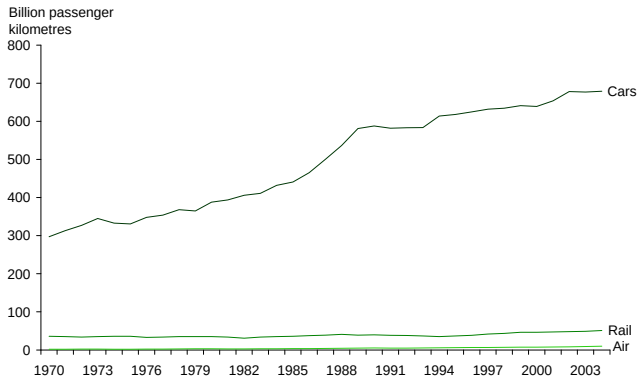
	1990	1995	2000
A	100	105	110
B	100	102	104
C	100	110	120

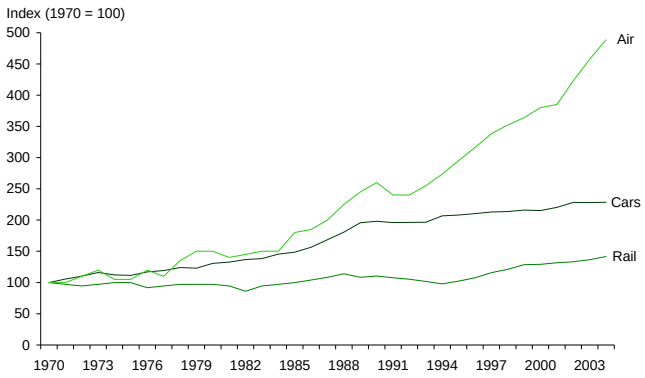
Here the relative trends are clear –C increased by the greatest proportionate amount, followed by A then B. Again we cannot say anything about their absolute values.

Why are they useful?

Index numbers are essential for **measuring changes in quantities that can not be directly measured**. The examples above also illustrate that producing indices for 'simple' data can be useful because of their simplicity, focus on change and their ability to **illustrate variations in series that would otherwise not be readily comparable**.

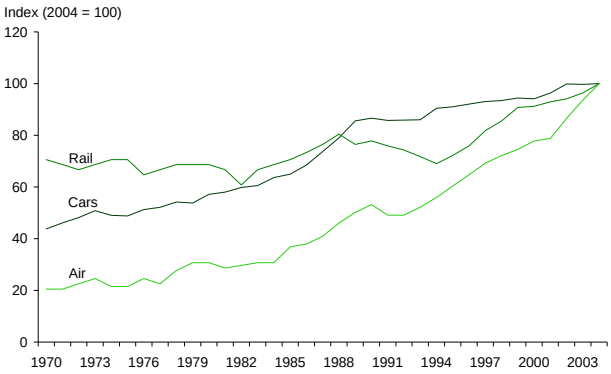
The charts below illustrates this, they look at trends in passenger transport by car, rail and air. The first chart tells us that transport by care was by far the most popular, but because it is so far ahead it is difficult to discern trends in rail or air travel. The second chart presents this in index form. The second shows very clearly the relative change of each mode over this period.





Because index numbers are based on changes they can often be **more up to date** than series with monetary values, for instance the average earnings index is updated monthly while the Annual Survey of Hours and Earning is, as its name suggests, only produced once a year. Similarly official statistics on energy prices and bills in monetary units are updated once a year, while monthly index values for energy prices are published as part of the RPI.

Potential problems



Problems can arise with interpreting index values if their values are **treated as actual numbers**. The guide on percentages contains more details. One problem specific to index number is the **choice of base year**. This can have affect how useful the series is and/or how it is interpreted. The reasons why 100 is most often used as a base value have been explained already, so it makes sense to make the most frequently used comparator the base figure. The chart opposite is based on the same data as the earlier chart, but with the final year set to 100. While it illustrates fundamentally the same numerical data, the patterns look different. It will generally be less useful because the base year is set at the end of the series, backward facing changes (1990 was x% less than 2004) are most easily made, while it is more intuitive to think of changes from an earlier to a later point (chronologically).

Other statistical literacy guides in this series:

- What is a billion? and other units
- How to understand and calculate percentages
- Index numbers
- Rounding and significant places
- Measures of average and spread
- How to read charts
- How to spot spin and inappropriate use of statistics
- A basic outline of samples and sampling
- Confidence intervals and statistical significance
- A basic outline of regression analysis
- Uncertainty and risk
- How to adjust for inflation