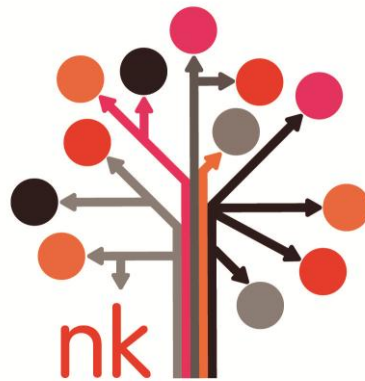# Northumberland Knowledge

Know Guide

Glossary of Terms

- November 2012 -

This page has been left blank

About this Guide

The Know Guides are a suite of documents that provide useful information about using data and information supplied via the Northumberland Knowledge website.

This Guide provides definitions of a range of terms used on the Northumberland Knowledge website.

For information on geographic terms used on the site see the Know Guide on geography.

Glossary of Terms

Average (mean)
The average (also known as the **mean**) is the sum of a set of numbers, divided by how many numbers there are in the set. So, for example, if the ages of five people are 26, 29, 31, 34 and 35, their mean age is (26+29+31+34+35)/5 = 31.

Census
A census is a complete count of a **population** (as opposed to a **sample**). In the UK a census of the entire population is taken every ten years - the last was taken in 2011.

Confidence Interval
A confidence interval is a range of values in which it is likely that the actual data value will occur. It is used to express the level of certainty that the estimated value accurately represents the true value for the population. The most commonly used confidence interval is 95 per cent. This means that, across the dataset as whole, the confidence intervals are expected to contain the true values around 95 per cent of the time. Confidence intervals are affected by the size of the sample used, so the smaller the sample size the less accurate estimates will be and the wider the confidence interval.

Dataset
A dataset is an organised collection of data that has a common theme.

Demographics
Demographics refer to the statistical study of a living population and particular groups within it.

Denominator/Numerator
The denominator is the total group that could have the characteristics being measured, and the numerator is the group known to have those characteristics. In calculations the numerator is divided by the denominator.

Gross Domestic Product (GDP)
Gross Domestic Product is a measure of economic activity. It is the sum of all incomes earned by the production of goods and services in the UK. It is equivalent to the value added to the economy by this activity. GDP is estimated in three ways: production (the value of goods and services produced),

income (the value of goods and services purchased) and expenditure (the value of income generated). In theory, these three methods should result in the same figure.

GVA (Gross Value Added)

Gross Value Added (GVA) is a measure of the value of goods and services produced in an industry or sector minus the costs of the raw materials or other inputs used to produce them. It can be described as the contribution to the economy (the GDP) of each individual producer, industry or sector.

Indices of Deprivation

The English Indices of Deprivation measure relative levels of deprivation in small areas of England called Lower Layer Super Output Areas (see Know Guide on geographic terms). They use 38 separate **indicators** covering seven distinct domains: income, employment, health, education, barriers to housing and services, environment and crime. These can be combined to attempt to calculate an overall measure of multiple deprivation, known as the Index of Multiple Deprivation.

Indicator

An indicator is a statistic that provides a summary of a specific theme or subject, eg crime or education and can be used to show change in that item over time or comparisons between areas.

Mean (see Average)

Median

The median is the middle value of a series of values when listed in size order. If the number of items is even the median is taken to be halfway between the middle pair of values.

Metadata

Document provided with a **dataset** which usually contains supporting information about the data such as; who supplied the data, how and where the data were collected, the variables in the dataset, and any other important supplementary notes.

Mid-Year Estimates

Annual estimates of the population based on the resident population on June 30th (the mid-year point) and calculated based on the census data but updated to allow for births, deaths and migration.

Mode

The mode is the most commonly occurring value within a dataset.

Numerator (see Denominator)

OAC

Output Area Classification (OAC) is a way of describing the characteristics of a small area. The Output Area Classification (OAC) distills key results from the Census for the whole of the UK to indicate the character of local areas. It profiles populations, structures other data, and helps target resources.

Percentile (see Quartile)

Population
A collection of all of the items about which we want to know some characteristics e.g. residents, employees or service users.

Projections
Population projections provide an indication of the size and age/sex structure of the future population if specified assumptions about future fertility, mortality and migration were to be realised.

Qualitative
Qualitative data cannot be measured numerically however it can be classified into categories. It can be used to help understand the characteristics of a population or an area. For example, model of car, type of house.

Quantitative
Quantitative data describes information that can be counted or measured numerically. It is essentially about measuring characteristics of a population or an area. For example, number of people or households.

Quartile (percentile)
A quartile is one of four equal groups that a population can be divided into. The lower quartile is the position below which 25% of the data falls while 75% of the data is below the upper quartile. The difference between the upper and lower quartile is called the inter-quartile range.

Range
The range is the difference between the lowest and highest value in a set of data.

Rate
A rate is a number or value expressed in relation to another number or value, for example, the number of burglaries per 1,000 households. It allows for comparisons to be made between areas or between data over time.

Sample
A sample is a sub-set of an overall population to be measured and is used to estimate the characteristics of a population. In many cases the point of a sample survey is to deduce information about the overall population and to do this the sample must be representative of the population.

Standard Deviation
The standard deviation is the square root of the **variance**. It is a measure of the spread of the data around the mean. It indicates how the data is distributed, e.g. is it clustered around the mean or more widely dispersed? Generally, if the standard deviation is small than values are clustered around the mean, if it is large then they are more widely dispersed.

Variance

The variance is a measure of dispersion of data indicating how far values lie from the mean. It is the average of the squared differences from the mean. Like the **standard deviation,** it shows how far the data is spread around the mean.

This page has been left blank

Northumberland Knowledge

The Know Guide was produced by the Policy and Research Team, Northumberland County Council
knowledge@northumberland.gov.uk

November 2012