Non-Technical Summary:
Practical Exemplars of the Analyses of Surveys (PEAS)

Overview

The main objective of the project was to develop a web site that would help people who are analysing social survey data to use the best statistical methods now available. The UK government, the Scottish Executive and the Economic and Social Research Council invest heavily in surveys to inform policies and to understand social processes. The anonymised data files from many surveys are available to policy makers within government and, free of charge, to researchers in UK Higher Education Institutions. The data are provided via the ESRC Data Archive and the Economic and Social Data Service supports users of the data.

The web site is built around a set of six exemplars each one using data from a different survey. The web site also has sections on the theory of survey design and analysis and on computer software that can be used to analyse surveys. There are extensive links between the sections so that, for example, the theory behind the use of a particular method can be accessed from the exemplar which uses it. The PEAS web site can be accessed at

http://www.napier.ac.uk/depts/fhls/peas/index.htm

Development of the Site

The material to go on the site was developed jointly by a survey methodologist, working for one of the largest social survey organisations in the UK, and an academic with an interest in the practical side of survey analysis. The site structure, architecture and design were engineered by a computer scientist and a web designer was employed to bring all of this together. Once enough material was available on the site, a series of three workshops were held with participants from academic institutions, government, local authorities and survey organisations. Feedback from the workshops informed the further development and improvement of the web site. In the final stages of the project email feedback, from experts and non-experts throughout the world was obtained via announcements on specialised mailing lists. This provided pointers to additional improvements and features that were included on the site.

Material covered on the site

Sometimes the respondents to a survey are not a good match for the population that the analyst wants to report on. This can happen deliberately if, for example, there has been a boost to the sample for certain types of respondent or inadvertently because it proves difficult to get some groups to respond. When this happens the survey analyst will usually need to carry out a weighted analysis that will bring the sample into line with the population. The PEAS site explains the correct way to carry out this type of analysis and illustrates how to calculate weights in different circumstances. The user can download an extract from the survey data from the site. Sample computer programs for four different types of software are available.

All results from surveys have margins of error associated with them. These margins of error depend on the sample size used in the survey with, bigger samples give smaller margins. But the margins of error can also depend on the sample design.

Some surveys use a clustered design, so that (for example) a sample of schoolchildren may consist of a set of clusters each of which is a whole class. This will usually increase the margins of error compared to what would have been obtained from a sample of children. If a survey is carefully designed so that the sample members are well balanced on other factors (e.g. the proportion of people living in different types of housing is balanced to match the proportions for the whole country) then the margins of error can become smaller. The exemplars help the analyst to understand how these features affect margins of error and how to allow for the survey design in the analysis using different software packages. Again, data and programs are available.

Missing data due to poor response rates and/or the failure of respondents to answer all the questions has become an increasing problem for survey methodologists in recent years. The PEAS site looks at different ways in which the survey results can be adjusted to make allowance for this non-response and exemplars, with tools to try them, are provided. The two commonly used methods are weighting, as discussed above, and imputation involving filling in the missing values. The PEAS site covers both of these and gives practical guidance as to how to attempt to deal with missing data. This is an area in which there is still considerable debate as to the best methodology to use. The site aims to take a balanced view by pointing out what is available but also suggesting ways in which analysts should look critically at their results to ensure that they appear valid.

The statistical software packages that are featured on the site (R, SAS, SPSS and STATA) were chosen because they are widely used by survey researchers. None of them is a specialist package that only does survey analysis.

Future developments

Software for survey analysis is a rapidly developing area. All four of the statistical packages had significant new features added during the lifetime of the PEAS project, some of which were prompted by our interaction with software providers. There will be a continuing need to keep the site up-to-date as further developments ensue. We also plan to continue to investigate procedures for dealing with missing data in collaboration with others.