Reorganisation of Dorset Councils: Interpretation of Results of Public Consultation

Now that the public consultation period has closed Opinion Research Services (ORS), independent analysts appointed by the Councils of Dorset, will be gathering together the various returns and preparing their report. They have provided a paper in which they outline how they will undertake their task. I have studied that paper and the purpose of this note is to outline the techniques to be used in the process they will adopt in non-technical terms.

The Sample Size

ORS will take into account all of the data arising from a number of sources: obviously the Consultation Questionnaire; the returns from the specially selected households across Dorset; input specifically from town and parish councils; the forums that they held across Dorset; other miscellaneous input such as letters received from residents.

The largest volume of data will arise from the Consultation Questionnaire and the returns from the specially selected households. Before we move on to say something about sample sizes I need to define two variables:

> The **margin of error** (sometimes known as the confidence interval) is a figure that defines a range of accuracy. So, when you read in the press of opinion polls that say 60% of the population favour Trident what they often don't tell you is the margin of error – but the statisticians know because they built it into their sample size. If the margin of error is 2% then what the opinion poll is really saying is that somewhere between 58% and 62% of the population favour Trident – but that caveat doesn't make for good headlines

> The **confidence level** tells you how sure you can be that 58% to 62% is correct. It is expressed as a percentage so a 95% confidence level means that 95 times in every 100 the statistical results are valid. Most researchers use a 95% confidence level as being sufficient for practical purposes. Again, this is rarely reported in the media.

If we apply this to Dorset and a population of 700,000 then to achieve a margin of error of 2% with a confidence level of 95% requires a statistical sample of about 2,400 people. The number of returns from both the Consultation Questionnaire and the specially selected households exceed that level which means the sample size is adequate.

A Representative Sample

That then raises the question of how representative the sample is. If I question 100 people picked at random in Saxon Square at 11.00 am on a Tuesday morning in August how likely is it that the opinions expressed will represent those of Christchurch residents? Not very, I would assess, because my sample will contain too many holidaymakers and not enough working residents.

Ideally, a selected sample is a miniature representation of the population it came from. Unfortunately, this is usually not the case in practice. One of the biggest problems is non-response, which may cause some groups to be either over or under-represented. Another common problem is self-selection where mainly pressure groups and others with a greater than usual level of interest participate more than other members of the population. When such problems occur, as is highly likely here, no reliable conclusions can be drawn from the observed survey data unless something has been done to correct for the skew.

A commonly applied correction technique is a **weighting adjustment**. It assigns an adjustment to each group of respondents. Persons in under-represented categories get a weight larger than 1 whilst those in over-represented groups get a weight smaller than 1.

By comparing the level of participation of a group with its known size in the overall population one can establish whether the survey response is representative with respect to this group. If there is a substantial difference between the response distribution and the population distribution, one can draw the conclusion that there is a skew in the survey data with respect to this group.

Suppose an online survey has been carried out. Among the variables measured is the age of respondents. Because the population distribution of age is known, one can compare the response distribution of age with the population distribution. *The following data is made up and is for illustrative purpose only.*

	Age up to 40	Age 41 to 60	Over 60
In population	30%	40%	30%
In survey sample	60%	30%	10%
Statistical weighting	30/60=0.5	40/30=1.3	30/10=3.0

Clearly those aged up to 40 are over-represented and those aged over 60 are under-represented. We can make the response representative with respect to age by assigning a weight equal to the distribution in the survey divided by the real distribution in the overall population – so for the over 60's a weight of 3.0

Thus in this *invented example* all data from the up to 40 group is halved in value, all that from the 41 to 60 group is increased by a third, whilst all data from the over 60 group is multiplied by 3 before it is carried forward into the main analysis.

Which People have responded

Acorn associated with the postcode will play a significant role in categorising responses. **Acorn** is a segmentation tool based around postcode which categorises the UK population into demographic types. It has been built by analysing social factors and population behaviour to provide information of the different types of people and communities across the UK. **Acorn** segments neighbourhoods into 6 categories, 18 groups and 62 types.

Acorn is used to categorise consumers' lifestyle, behaviour and attitudes, together with the needs of communities and is important to both private sector and public service organisations. It is used to evaluate local areas and to identify the characteristics of each neighbourhood.

In Summary

I would expect the selected households' survey to be reasonably sound because the households were selected to represent the demographics of Dorset - provided responses are spread across Dorset and not predominantly from one or two areas. I would expect ORS to comment upon this aspect of the data.

The Consultation Questionnaire will have its own profile of respondents but that is unlikely to match the profile of Dorset as a whole. Most of the Questionnaires were filed on-line and that alone will potentially skew the sample. Weighting will almost certainly be required on this data and ORS should tell us how they have done that.

ORS will prepare an analysis of both sources of data and will in addition compare and contrast the two.

There will then be a further report on the other sources of data. It is important that ORS have stated that they will not simply attempt to lump all these findings together. Such an approach would be quite wrong because volunteered opinion (in letters say) and different methods of extracting data from people (focus groups say) produce different results and it is also important that the views emanating from any particular region are viewed together and not lost by being mixed in with views from a different area.

A vital consideration in undertaking this type of analysis is that a consultation is neither an election nor a referendum. This is not simply a numbers game. The analyst is expected to discover trend, to highlight issues and anomaly, to summarise various aspects – all aimed at assisting Councillors to reach rational conclusions that have a foundation in public opinion. OPS make it very clear that they are fully cognisant of that aspect of their function.

Jim Biggin

October 2016