Summer 2014

be selective



# Select News...

## Predicting Energy Efficient Households

Select has been working with the Energy Saving Trust (EST) to develop their Home Analytics product for Scotland. This product is designed to assist the Scottish Government and Local Authorities to deliver energy efficiency retrofit measures such as



cavity wall insulation in properties across the country.

Home Analytics is an address-level data set for the housing stock in Scotland, providing information on property characteristics and the potential for energy efficiency retrofit measures, as well as other sociodemographic information. This data set is critical for Local Authorities since it allows them to target and provide funding to those properties that have not yet benefited from retrofit measures.

Home Analytics is based upon a variety of data sources, such as Energy Performance Certificates and boiler installation records but, even combined, these sources are unable to provide a complete record of all Scottish properties. We therefore worked with the EST to build a statistical model that could be used to predict property characteristics and fill in these missing gaps.

The statistical model was developed so that it used the characteristics of the surrounding properties to assess the probability that a property is, for example, terraced, a flat, semi-detached or detached. The model assigned a probability to each category so that the missing record could be replaced by the category with the highest probability, with the probabilities themselves used to provide aggregated summaries at different spatial scales.

By filling in the missing records in Home Analytics, the Scottish Government is better equiped to identify houses in need of energy efficiency measures and to ensure that funding is directed at those most in need.

#### **Our Services:**

#### Advice

Our statistical advice service offers detailed expert guidance on all areas of statistical analysis and interpretation to allow your staff to understand or perform their own statistical investigations.

#### Analysis

Our statistical analysis service aims to identify and apply the most appropriate statistical tools to make best use of your data in addressing the questions that you need answering.



#### Research

Our statistical research service extends the analysis service above to tackle problems that lie beyond the boundaries of current statistical know-how perhaps because of the inherent complexity of your data or simply because the usual underlying assumptions don't apply.

#### Training

We provide training for statisticians and nonstatisticians alike and can cover any statistical topic from the applied end right through to cuttingedge methodological research.



#### Data

We can source, collect, collate and cleanse data either simply for presentation direct to the client or for subsequent analysis.

#### Surveys

Often the easiest and best method to obtain information you need from people is simply to ask them. We can conduct online, phone and face to face surveys starting from initial design through to final analysis.

#### Making numbers work for you

If you would like more information on any of the Services that Select Statistics offer, please contact us on 01392 440426, email: info@select-statistics.co.uk or complete the online contact form on our website: www.select-statistics.co.uk



## Select's Free Online Calculators

We've recently been adding to our website to include some online statistical calculators to help you carry out some basic statistical calculations of your own.

To start off this new resources section, we've chosen to add calculators on the types of analysis that we commonly undertake.

The first set of calculators are for standard sample size calculations. With a small number of inputs (for example, information on your study design such as the population size or the confidence level you require), you can obtain the sample size for your study.

Confidence interval calculators are also included so that you can quantify the uncertainty associated with an estimate of, for example, a population mean. The confidence interval is the range of values in which we estimate the point estimate to lie given our level of confidence.

You can access our calculators via the Resources tab on our website or via the weblink: http://bit.ly/ PKPNpF.

We're planning on adding to these online calculators to include other types of statistical analyses such as statistical hypothesis testing. If you have suggestions for new calculators that you'd like us to create, we'd be pleased to hear from you! You can contact us via email: info@ select-statistics.co.uk or phone: 01392 440426.

## Testing the Effectiveness of Treatments

Select has been working with Unex Designs Ltd. to test the effectiveness of their Pet Remedy treatment for dogs with behavioural issues. We helped with the design and analysis of a new dog trial for Pet Remedy, comparing the effectiveness of the treatment vs. placebo, both administered in combination with behavioural therapy.

We advised on the study design, helping to identify and limit potential sources of bias from the trial results and to ensure a fair comparison between the treatment groups.

The design included randomisation to the treatment groups to help protect against systematic differences at the start of the experiment. Double-blinding was also recommended, with treatment information being withheld from both the dog owners/handlers and trainers/assessors, to help ensure that (even subconsciously) dogs in the different treatment groups were not handled differently, which might also have had a biasing effect on the results.

Sample size calculations were also used to estimate the likely number of dogs required to be able to detect a statistically significant difference between the treatments in improving the dogs' behaviour.

The study took place over a three month period of time with monthly observations recorded that reflected the dogs' behaviour and excitability. Measurements were also taken before the study to establish the behaviour and excitement levels of a dog pre-treatment. This allows us to assess the progress of the dogs compared to how they behaved at the start of the trial.



In the end, the study showed a general improvement in the behaviour of all dogs during the study period, suggesting that the behavioural therapy was an effective treatment in its own right. However, a statistically significant difference was found between the treatments showing that dogs

on Pet Remedy improved significantly more than those on the placebo.

This was great news for Unex as this study provides them with more evidence that Pet Remedy is an effective means of dealing with dogs with behavioural problems above and beyond the improvement observed when behavioural therapy is used on its own.

#### Select presents the RSS Challenge 2014

We are pleased to be sponsoring the Royal Statistical Society's Statistical Analytics Challenge 2014. Over the Summer, teams of analysts from across the country will be testing their analytic skills, teamwork and creativity with the aim of being named the 2014 RSS Challenge Winners! The brief, unveiled on the 6th May, is an open-ended research question that relates to a large and complex data set. This year, the challenge involves analysing brain imaging data from MRI scans. Due to its complex structure and the huge amounts of data that MRI studies amass, statistics plays a crucial role in identifying key insights from these data sets. The RSS will judge the entries and choose the top three teams from the challenge who will be invited to present their results in a special session at the RSS Conference in September. More information about the challenge is available at the following link: http://rsschallenge.wordpress.com/.

## Putting data to work

## Case Study: Forecasting Household Waste for Essex County Council

County Councils are working towards developing cost-efficient and sustainable waste management plans so that landfill is the option of last resort with preference given to prevention, reuse, recycling, composting and energy recovery alternatives. In order to develop these plans, Select worked with Essex County Council to develop short to medium term forecasts of waste that they could then use to assess their future operational needs.

Changes in waste tonnage and composition are likely to be influenced by a number of transient and long-term drivers including: increasing population; changes in demographics and household occupancy rates; changes in public attitudes towards waste generation and recycling; changes in consumer demands; as well as economic factors.

By collecting and collating local council-level data on the volume of household waste collected as well as the composition of waste over a period of years, together with data on potential drivers such as those listed above, we were able to identify the most important factors affecting the volume and composition of waste produced. These data were combined within a statistical model to predict the waste tonnage in future years, taking account of past trends and incorporating future projections of the relevant driving factors.

The model was developed so that it could be implemented within a tool (for example, Microsoft Excel) allowing Essex County Council to predict future waste tonnages under different scenarios in each region and to break these down by different waste types. These projections of municipal waste provide the council with a basis to assess future operational needs and to develop cost-efficient and sustainable waste management plans.



#### Interested in more case studies?

Check out our website for more case studies in marketing, retail and many other areas, as well as details on the many statistical services that Select offers.

## **ExIStA Update**

The Exeter Initiative for Statistics and its Applications (ExIStA) has had a great start to the year. In January we hosted a workshop including three diverse and fascinating talks on building a tennis model in 15 minutes, climate modelling and how statistics can be used to improve working and living environments. There was a great turnout to the event with plenty of questions and discussions over tea and cake after the talks.

Following the success of the previous joint ExIStA/ RSS South West Local Group Meeting, we were pleased to have invited speaker Professor Adrian Raftery join us in February to talk about his work in probabilistic weather forecasting. The topic of this seminar was particularly relevant for both the the Exeter Climate Systems group at the University of Exeter and researchers at the Met Office.

In addition to these ExIStA events, 2014 has seen the launch of the Exeter R Users Group, a group setup for anyone with an interest in the statistical language R, whether they are expert users or just starting out. The launch happened on 27th February with a variety of speakers talking about how they use R in their statistical analysis followed by networking. The event was a great success with over 75 attendees.

If you'd like to find out more about ExIStA, including details of our upcoming events visit our website: www.exista.org or contact us on events@exista.org.

#### Sarah is awarded her Chartered Status

We're pleased to announce that Sarah has been awarded the prestigious status of Chartered Statistician by the Royal Statistical Society, recognising her extensive training and experience as a professional statistician.

The Chartered Statistician (CStat) is the RSS's highest professional award providing formal recognition of an individual's work experience, formal qualifications and professional training. To qualify, the RSS require an approved degree together with at least 5 years' post-graduate training as a professional statistician.

Sarah's award recognises her Master's degree training together with four years' experience at the University of Sheffield Statistical Services Unit and 18 months as a statistical consultant here at Select. She also demonstrated a strong and consistent commitment to continued professional development (CPD).



## Focus on... Business Analytics

Would you like to be able to gain a competitive advantage and learn more about your business and customers?

Business Analytics is about using your data to provide meaningful business information so that you can make better decisions, develop strategy or improve performance. It can relate to your company's products, services, processes and staff to ensure that they are all working efficiently and effectively.

Simple techniques can be used to answer questions such as:

- Who are your customers and what are their preferences?
- What do your customers buy and at what price?
- Do you have sufficient stock and cash flow?
- How can you allocate your staff to minimise waiting times?
- Which web layout is the most effective?

Here are just three examples of how we can help:

- Using data about your customer base, we can assign customers into groups with similar preferences or characteristics. Understanding these groups means that you can adopt a different marketing mix per group, target those groups with the best prospects or identify the groups with the highest likely lifetime value.
- Testing a marketing campaign or new website layout on a sample of customers would allow us to assess their effectiveness before rolling them out to your entire customer base. We can also use this information to identify the attributes of customers that are more likely to convert given a campaign or layout.
- 3. We can combine your business data with other sources to identify the key factors that drive stock demand such as advertising or the weather. Combining these factors in a tool to forecast stock allows you to be confident that you have sufficient inventory to meet customer demand and assess how and when to place new orders.

To find out how we can help you analyse your business data call us on 01392 440426.

## The Select Network

Select is supported by an extensive network of external consultants based at the UK's top Universities and research institutes and also by a community of specially-selected partner organisations offering skills and services that are likely to be of value to our clients. The Select Network provides us with an enhanced range of skills and expertise to deliver service and value to our clients. Details on how to register your interest in the network are provided in the box at the bottom of the page.



#### Meet Henry Burroughs



Henry is a statistical consultant specialising in the analysis of survey and operational data to inform business decisions. He holds a BSc in Mathematics from the University of Leeds and a Master's degree in Statistics from University College London. His statistical

expertise includes cluster analysis for market segmentation, logit modelling for price sensitivity and multivariate analysis.

Henry has delivered several high profile social and market research studies involving a variety of statistical analysis techniques and interpretation to national and local government, regulatory bodies and businesses. His career experience ranges from researching cancer epidemiology, measuring travel behaviour and demand, to measuring the impact of marketing communication campaigns for London 2012.

#### Would you like to join the growing Select Network?

If you are interested in joining the Select Network as a freelance consultant or would like more information, please contact us on: 01392 440426, email: info@select-statistics. co.uk or complete the online contact form on our website: www.select-statistics.co.uk

