

Statistics: For Businesses, About Businesses

A one-day seminar about collecting and producing official statistics on businesses and showcasing potential uses in business, for research and for policy making, organised by the Quality Improvement and Official Statistics Sections of the RSS and ENBES (the European Network for Better Establishment Statistics)

Tuesday 19th February 2013 10am to 5pm

at the

Royal Statistical Society, Errol Street, London, EC1Y 8LX

<http://www.rss.org.uk/site/cms/contentCategoryView.asp?category=142>

The seminar aims to highlight actual topics and issues in collecting and producing official statistics on businesses and, on the other hand, showcase potential uses of official statistics and businesses own data for business, for research or for policy making. In spite of the vast importance of business statistics for policy making at all levels of society, this branch of statistics has received comparatively less research and methodological attention.

Since its initiation in 2008, the European Network for Better Establishment Statistics (ENBES) has worked on advancing exchange between practitioners, methodologists, and academics on matters relating to business statistics. In parallel with - amongst others - the European Commission's MEETS programme and the FP7 research project BLUE-ETS that were initiated at approximately the same time, ENBES is endeavouring to bring business statistics closer to its users, as well as helping understand user needs for business statistics better.

Therefore, in addition to its biennial European Establishment Statistics Workshop series, ENBES is launching regional one-day seminars with the intention of bringing together practitioners, users and methodologists in business statistics to improve mutual understanding of perspectives on topics and issues, and stimulate more effective use by businesses of official business statistics data.

The first seminar in the series, co-organised with the Official Statistics Section (OSS) and Quality Improvement Section (QIS) of the Royal Statistical Society (RSS), and to be held at the RSS in London on February 19, 2013, will offer a balanced set of up-to-date presentations including an overview of the field, case studies of data use and issues arising therein, importance and potential of new data sources, as well as a summarising keynote talk.

Agenda

10.00-10.10 Introduction – Boris Lorenc (ENBES)

10.10-10.50 Sampling and estimation procedures in business surveys: a discussion of some specific features – David Haziza (Université de Montréal and CREST/ENSAI)

Business surveys present important methodological challenges. Many are discussed in the excellent paper of Rivière (2002). In the presentation, we will discuss some specific aspects of sampling and estimation encountered in business surveys: (i) sample coordination; (ii) the treatment of nonresponse including the construction of weighting and imputation classes and the problem of composite imputation and (iii) the treatment of influential units. For (iii), we will define the concept of influential values, present the conditional bias of a unit as a measure of influence and discuss the properties of winsorized estimators.

10.50-11.20 Coffee break

11.20-12.00 Survey and Statistical Communication – Jacqui Jones and Kate Davies (Office for National Statistics)

Communication is one of the central elements for maximising the quality and utility of statistics. From a survey perspective communication is essential so respondents understand why their data are being requested and how it will be used i.e. what statistics will be produced from their data. From a dissemination perspective, communication is essential so people are enticed to look at the disseminated statistics, understand what the statistics mean and any limitations to their use.

For survey respondents, the quality of survey communication is often associated with the level of attention that they give to the survey request and ultimately the quality of data they return. In turn this will impact on the quality of the derived statistics and respondents perceptions of burden.

For statistical users, the quality and utility of communication is often associated not only with the statistical commentary and presentation but also the types of products produced.

Over the past few years work has been undertaken at the Office for National Statistics to improve survey communication, the provision of information about the quality of the statistics, and more recently work to improve statistical commentary. Work has also begun to look at the range of statistical products produced and move away from a one product fits all approach, so we produce products to meet the needs of a wide range of users e.g. statistical connoisseurs, less frequent users and new users.

The objective of this paper is to provide an overview of current and possible initiatives to improve survey and statistical communication at ONS.

12.00-12.40 Intelligent data analysis to aid decision making in a commercial environment – Gordon Blunt (Gordon Blunt Analytics Ltd.)

With the power of modern hardware and software, many companies nowadays possess large databases. Given that there is a computer on virtually every desktop, it should be easier than ever for companies to make good use of their data. However, there are many pitfalls on the road to insight, such as the quality of the data, the use of operational systems that were never designed for analysis, legacy systems that do not fit well with modern software, missing or incomplete records and so on and so on.

One of the most common pieces of business software is the spreadsheet. Spreadsheets were first introduced in the 1970s, and, in their early days, had limited data handling capabilities. Given the ubiquity of modern desktop computers and the powerful spreadsheet software now available, anyone who has a data set can produce their own analysis. This 'democratisation of analysis' can give statisticians many opportunities, but it may also lead to problems.

Making sense of data has always needed appropriate skills and knowledge, regardless of the size of the data set. Further, statisticians have always known the importance of examining their data, but there are particular problems with large data sets, and visualisation is critical – both for examining structures and identifying data problems. Errors in a large data set may not be trivial to correct; indeed, the larger the data set, the harder the task.

This talk will suggest ways to address these – apparently – contradictory elements to ensure that companies can make the best use of data. Use of external data to help develop insight will be described too, because such data may be the only way a company knows how it is performing compared to its competitors.

12.40 Lunch break (lunch is included in the registration cost)

14.00-14.20 Open Data Institute - Stuart Coleman (Commercial Director, Open Data Institute)

An introduction to the ODI's mission and progress to date

14.20-14.50 Case study of business use – Nigel Marriott (Marriott Statistical Consulting Ltd)

Innovantage collate nearly all on-line job adverts using a proprietary web search system. This results in a database consisting of approximately 1.5 million job adverts from nearly 200 job boards every month which makes it an extremely rich source of intelligence to anyone interested in labour market trends in the UK at both the macro & micro level. A major barrier to providing high quality market intelligence is that a single job advert can be posted on multiple job boards, by multiple recruitment agencies and reposted multiple times by job boards to drive up their apparent traffic. Stripping out all of this duplication in order to arrive the true number of job vacancies is a significant statistical exercise which has been undertaken by Nigel Marriott of Marriott Statistical Consulting Ltd.

Crucial to the task of identifying of whether the correct deduplication model had been built was the job vacancies data supplied by the ONS using their Vacancy Survey and Labour Force survey. In his talk, Nigel will talk about how the ONS data was used to develop new methods of deduplication, identify unexpected data quality issues and to inspire a radical way of overcoming a barrier to labour market intelligence that was thought to be insurmountable by all participants in the industry.

14.50-15.30 Extracting the signal for policy from the noise of statistics – Christopher Hackworth (Bank of England)

Policy makers have a range of official and private sector statistics at their disposal when considering the outlook for the economy, and their policy decisions. This presentation will discuss some of the tactics and methods some Bank staff use to help form judgements about what has been going on in the economy over the past, and come to a near term view of where the economy is likely to go.

15.30-16.00 Tea break

16.00-16.30 British business statistics - public sector ambivalence and private sector opportunity – Stephen Penneck (President of the International Association for Official Statistics and former Director General of the Office for National Statistics)

16.30-16.50 Discussion

16.50-17.00 Closing remarks