Australian Government Statistical Forum

Wednesday 13 October 2004
2.00pm - 5.00pm

Agenda & Papers

Australian Bureau of Agricultural and Resource Economics
The Edmund Barton Building,
Core 6
Corner of Broughton and Macquarie Sts
Barton, ACT

1. *Minutes* from previous meeting, 13 May 2004 at Australian Institute of Health and Welfare.

2. *National Statistical Service*

(i) National Data Network (NDN) - The NDN is an initiative which will provide infrastructure, protocols, standards and services to support the sharing and integration of data across Australia. Jonathan Palmer from the ABS will provide an update on recent developments.

(ii) Finding Information: NSS Search Portal - This paper provides details on the search facility that will be released on the NSS web site for users to search across a number of statistical directories.

3. *Strategic Review of ABS*
Dennis Trewin will provide information on progress with a review of the ABS, which focuses on future strategic directions.

4. *Disseminating Data*
Three case studies on disseminating data will be presented by ABARE, Centrelink and HIC. The presentations will cover dissemination tools, developments over the last few years and what is planned for the future.

5. *Why Statistical Classifications and Standards make a difference?*
The ABS will make a presentation on how it develops classifications and standards. It will look at examples where they made a positive difference and where the failure to use them led to problems.

6. *Expanding Australia's Census*
The next Population Census provides an opportunity to establish the basis of a longitudinal data set from successive Censuses, and/or to link Census and other ABS data sets. This paper outlines potential uses of linked datasets and foreshadows initial testing that will be required. A verbal update on progress will be provided.

7. *Information Development Plans*
This paper provides an introduction to Information Development Plans, which aim to improve the available information for a particular field of statistics. Three fields will be explored, Agricultural Statistics, Water Statistics and Rural and Regional Statistics.

8. *International Statistical Activities*
Round table discussion on international statistical meetings attended by members.

9. Statistical Training
The ABS will provide information on consultancy services and training courses currently available to Australian Government agencies.

10. Arrangements for next meeting

Finding Information: National Statistical Service (NSS) Search Portal

Background

1.   One of the objectives of the NSS is to increase the availability, accessibility and useability of information derived from key administrative and survey data sets to provide the community with a wider range of statistics to assist decision making. The NSS web site was released in February 2004 to support these objectives. The site contains material on sound practice for the collection and use of information. To facilitate this objective the NSS web site provides links to the statistical directories of a number of government agencies at the Australian and State Government levels. These links provide a useful first step in allowing users to search for relevant data.

2.   Long term, the NDN should provide a comprehensive suite of programs to assist producers and provide a single entry point for users to search for data. However, in the short term, a search facility has been developed to maximise the accessibility of the metadata currently available through the web.

NSS Search Portal

3.   The NSS Search Portal will provide a basic search facility across a range of government web sites for users to find statistical metadata. Initially, the search will be a basic document search. This means that when you enter a search word or phrase, the facility will return documents containing that word or phrase. Users can also limit the search to a specific agency.

4.   The NSS Search Portal 'spiders' (similar to yahoo and google) through web sites using the metadata standard Australian Government Locator Service (AGLS). It uses a combination of items to build its index including the title and body of a page. For the search facility to operate effectively all pages on relevant web sites need to be indexed.

5.   The AGLS Metadata Standards are based on the Dublin Core Standard, which is a set of 19 descriptive elements which government departments and agencies (including states and territories) can use to improve the visibility and accessibility of their services and information over the Internet. These elements include a description of the page contents as well as the title, creator, date, topic etc.

Sites which are currently Searched

6.   The NSS Search facility will enable you to find PDF or HTML documents or URLs that match your search criteria. Results will be taken from the research and statistics areas of nominated Australian Government web sites that are AGLS compliant. Initially, the search facility will cover the following web sites:

- Office of Economic and Statistical Research, Queensland
- Australian Government Department of Agriculture, Fisheries and Forestry
- Australian Institute of Health and Welfare
- Australian Bureau of Statistics
- Statistical Clearing House

7.   This list will be expanded as more sites are added and as more agencies incorporate AGLS tags into their web sites.

How can agencies be included in the Search Portal?

8.   The Australian Bureau of Statistics welcomes agencies to include their statistical directories in the search facility on the NSS web site as the effectiveness of the facility relies on the participation of government agencies. We are willing to assist agencies in meta-tagging of the statistical components of their sites. If agencies only have a small number of collections, they can have their metadata loaded into a Caretaker's Directory located on the Statistical Clearing House web site.

9.   For more information or for assistance in applying the AGLS tags, please email inquiries@nss.gov.au or ring the NSS information line 02 6252 5785.

EXPANDING AUSTRALIA'S CENSUS

Issues

> The ABS is currently assessing the possibility, from the 2006 Census onwards, of establishing a longitudinal data set from successive censuses, and/or linking census and other ABS data sets.
>
> ABS is currently testing public reaction to data linking. ABS will not proceed if there is a risk that it may jeopardise the success of the census.
>
> This paper outlines the analyses that such data sets might support, in a policy context.

Known Facts

> There is user interest in analysing transitions and outcomes, which are best supported by longitudinal data sets. Linking Population Census data longitudinally or with other data sets would potentially provide a rich data set to support a range of analyses, and provide the basis for improvement in community outcomes.

Conclusions/Options

> This paper explores a range of potential uses of linked data.

Expanding Australia's Census

Expanding Australia's Census

1. Background

1.1   An ABS Information Paper setting out preliminary ABS Views on Content and Procedures 2006 (Cat. No. 2007.0) [1] was released on 2 July 2003. It included a discussion on improving the useability of census data by combining them with data from other sources, including ABS surveys. (p10; see Attachment 1)

1.2   The information paper invited submissions from users of census data and members of the public on this or other issues discussed in the information paper. Four submissions were received on this issue, all supporting ABS exploring means to improve the useability of census data, from

- Department of Agriculture Fisheries and Forestry,
- Commonwealth Grants Commission,
- Australian Institute of Health and Welfare, and
- National Health Information Management Group.

1.3   These Submissions discussed the potential benefits of better relating population and housing census data to

- Data from the Agricultural census
- Both other ABS surveys, and administrative data sets, to bring together information for population groups such as the Indigenous population to compile group specific socioeconomic indexes
- Health related data from administrative sources, including registers of death or incidence of disease.

1.4   A further two submissions were received seeking changes to the census questionnaire to enable better identification of farmers and farm households. Their statistical objective would be met by linking of Population Census and Agricultural Census records.

1.5   No submissions were received arguing against further investigation of the potential to improve useability of census data.

1.6   At its May 2004 meeting, the Australian Statistics Advisory Council was strongly supportive of ABS undertaking further planning for linkage of population census data, commencing with the 2006 Census, and doing further community sensitivity testing.

1.7   The purpose of this paper is to outline some potential instances where there could be significant enhancement of the value of census data, and to set those instances in a policy-relevant context. The examples are illustrative rather than exhaustive of the potential for expanding Australia's Census of Population and Housing.

1.8   While the 5-yearly Censuses, and other data sets referred to in this paper, support

considerable cross-sectional analysis on many of the social issues and population groups discussed below, they are unable to support adequately the analysis of longitudinal perspectives, such as transitions and pathways over the years. There has been a growing demand for these perspectives, which has become evident in the commissioning and funding of various longitudinal surveys.

1.9   ABS is currently assessing public reaction to data linking and investigating possible protocols that might be established to protect confidentiality of the data and maintain the confidence of the public. A final decision whether to go ahead is expected by May 2005.

1.10   Public reaction to the following three options will be tested:

1. Linking of census data with data collected in successive censuses, other ABS run collections and with Births and Death data,
2. as above, plus linking of census data with health data;
3. as above, plus linking of census data with other administrative datasets collected under the *Census and Statistics Act*.

1.11   The following points should be noted:

1. The ABS will not proceed with any linking exercise if there is a risk that it may jeopardise the success of the census. That is, testing must clearly show that there is community support for such linkage, and that respondents will still complete the census form if data linking is to proceed. Both qualitative and quantitative testing will be undertaken to determine the potential impact of data linking on respondents and Census response rates.

2. Record linking will not be undertaken using data collected prior to the 2006 Census. This is because the ABS has not informed respondents of such use of past data at the time it was collected. It is also impractical as information needed for record linking is not available.

3. The ABS will need to develop a fuller understanding of the relative benefits and risks of including all people in such a linkage exercise, or just a sample. Both options would require the retention of a name based match key in association with the unit level data.

4. Only datasets collected under the *Census and Statistics Act* will be considered for linking. This could be data collected directly by the ABS or it could consist of other datasets collected by other agencies and brought into the ABS under the Act.

5. As required by the Census and Statistics Act, no identifiable data, linked or otherwise, will be released from the ABS. This is likely to impose quite tight constraints on the information that can be made available through CURFS. However it is expected significant analyses of microdata will be possible through CURFs, through RADL and through other data lab approaches.

2. Types of Data Linking

2.1   There are different types of record linkage, and the techniques used depend upon both the availability of information to make the links, and the intended purpose of the combined data set. Bernier and Nobrega [9] state

"there are two categories of record linkage: exact matching and statistical matching. The purpose of an exact match is to link information about a particular record on one file to information on a secondary file, thus creating a single file with correct information at the level of the record. This is achieved by using deterministic or probabilistic record linkage procedures. However, the purpose of a statistical match is to create a file reflecting the current population distribution, records that are combined do not necessarily correspond to the same physical entity (e.g. person, business or farm). Statistical matching involves modelling and imputation techniques to complete a file"

2.2   This paper is focussed on potential uses of composite data sets formed by exact matching. Bernier and Nobrega further state that:

"an exact match is when files are linked at the level of the record. Exact matching can be either deterministic or probabilistic. .....In order to perform a deterministic exact match, the variable must be complete and exactly the same on both files....... Probabilistic record linkage is another form of exact matching...... (it) can compensate if the information is incomplete and/or subject to error"

This is the likely situation for most of the opportunities for linkage discussed in this paper.

3. Sources Which Could Enrich Analysis of Census Data

3.1   If it is decided to proceed with a strategy to expand Census data, commencing with the 2006 Census of Population and Housing, there are several constraints to what would be done. The ABS has decided that the scope of any linking would be only to data sets held by the ABS now and in the future. All linking would occur within the ABS and the linked data set would be subject to the same confidentiality restrictions as other ABS data sets.

3.2   Within these constraints, options for combining data from various data sets include

(a) Censuses only

(1) For each successive census, starting from 2006, compile a longitudinal data set of individuals for the full population, for all census variables.

(2) As for (1), but for a selected set of variables of major interest for longitudinal analysis.

(3) As for (1), but for a sample of the population, rather than the full population.

(4) A reduced data set for a sample of the population.

(b) Censuses and other ABS Household Surveys

Censuses, of necessity, have a quite limited set of questions on any specific topic. However, because those questions are asked in relation to everyone, the Census can support small area analyses, and analyses relating to quite small population subgroups. The various household surveys, however, contain much more detail about the particular topics. Composite data sets derived from both Censuses and sample surveys, could support more extensive analysis than is possible at present.

(1) For all respondents in specific future household surveys, such as the National Health Survey, the Survey of Disability Aging and Carers, the General Social Survey, the Household Expenditure Survey, the Survey of Education and Training, etc., include relevant Population Census data items from 2006 and future Censuses.

(2) For all census respondents, add to the data set relevant variables from future ABS household surveys, as individuals are selected in them.

(c) Censuses and other ABS Surveys

(1) There is policy interest in having data from the Population Census and the Agriculture Census combined into a data set to support analysis of the economic and social aspects of Australia's farming community.

(2) Similarly, economic performance data for owner/operator and "family" businesses could be related to social aspects for very small businesses.

(3) Transport analysts who use data on motor vehicle usage would be interested in the social characteristics of owners/households where vehicles are garaged.

(d) Censuses and Administrative Data Sets held by ABS

(1) ABS uses data collected as a by-product of administrative processes. Examples are the vital registrations data, and characteristics of persons recorded as committing crimes, the subject of criminal court cases, and sentenced to spend time in a correctional facility.

(e) Census and Administrative Data Sets not currently held by ABS but which could be provided to ABS

Many government agencies hold data relating to persons to whom services are provided or with whom they interact. While in most cases this data is not currently available to the ABS, a case for ABS access might be made where there was a strong statistical benefit in having such data combined in some way with census data for analysis. (In some cases, there would need to be legislative or administrative action to allow this to occur). Examples of such data sets include:

- Centrelink payments under various programs which it administers
- Data from the FACS Longitudinal Data Warehouse
- Data from the FACS/DEWR Joint Data Set and Working Age Longitudinal

Extract
- Personal income tax data from ATO
- Details of services for which Medicare or Pharmaceutical Benefit payments are made by HIC
- Details of various categories of immigrants admitted to reside in Australia, held by DIMIA
- Details of persons admitted to hospitals, or diagnosed with specified illnesses (hospital morbidity and disease register data held by AIHW).
- Details of persons receiving support for cost of housing.

4. Social Issues and Population Groups.

4.1   Social issues, policy, and programs often are multidimensional and complex in nature, and in terms of the people involved. Social statistics similarly need to be multidimensional to support the analysis, evaluation and policy formulation to address those issues.

4.2   There are two basic dimensions of social statistics which are useful for addressing the potential ways in which population census data could be expanded - areas of social concern, and selected population groups.

4.3   The ABS [2] has identified eight major **areas of concern:**

- Family and community
- Health
- Education and Training
- Work
- Economic resources
- Housing
- Crime and justice
- Culture and leisure

These areas of social concern are neither mutually exclusive nor exhaustive, but they provide a useful framework and focal points for developing social statistics.

4.4   This focus is further sharpened by relating these areas to the general **population groups** which the ABS has identified as of particular interest. Examples are:

- Unemployed people
- Retirees
- Indigenous people
- Lone parents
- Children
- Migrants
- Older people
- People with low income
- People with disabilities
- Crime victims

(Again there is a degree of overlap between these population groups; typically a person may belong to several of these groups).

4.5   Other policy priorities that have arisen and which need data to better inform future discussion and initiatives are

- Geographic dimensions - rural and regional community needs, both economic and social; impact of people on the environment

- Intergenerational issues, including future population and labour force size and composition social stress and capacity to support economic growth and meeting future demands for government provided services.

5. Family and Community

5.1   There has been a major change in recent decades in the structures, functioning and levels of dysfunction in Australia's families. There have been several factors. Analysis of the impacts of changing living and family arrangements, and development of policy to address social concerns, and particularly areas of social disadvantage, has been limited by the scarcity of data that trace outcomes. Recent initiatives to conduct longitudinal surveys of income and labour dynamics and associated topics, and children, will begin to address this in a limited way, although such longitudinal surveys are very costly, have limited sample size, and may be constrained by falling response rates and biases in results.

5.2   Development of data sets derived from successive population censuses would provide a basis for analysis of issues such as

- Differential education completion rates and labour force entry and experience for children growing up in different household types.

- Geographic mobility and type of housing

5.3   Because of the various support programs available for some families, further extension of such a data set to include information on entitlement to and receipt of such benefits could be contemplated, either by asking respondents in a household survey conducted by ABS, or by augmenting the census data with administrative by-product data extracted from Centrelink data sets.

5.4   Issues related to care, within the family unit and community and more broadly, are changing and assuming greater policy importance. Information about care for children, older people and people with disability are needed to analyse impacts of present arrangements and support development of policies and programs into the future. The 2006 Population Census is likely to include questions on disability and on provision of unpaid care for persons with disability, long term illness or old age, and for children. While these questions will provide very valuable data in respect of 2006, the potential to assess these longitudinally with data from future censuses to understand the duration, impacts and consequences over time would increase potential value and support the evaluation of disability in a broader social context. Longitudinal analysis will give richer

insights than the cross-sectional analysis, largely because it will support a greater understanding of impacts on individuals and families over time. Policy issues to be addressed are:

- Is the carer population relatively constant over time, or does it change?
- What are the characteristics of carers? Are they combining caring with work and/or education and training? How do these other commitments change over time?
- When people become carers, what was their former labour force status? Has it changed?
- When people cease to be carers, to what extent do they enter the labour force or increase time at work?
- Are there differences between metropolitan, non metropolitan urban, and rural areas in the provision of care? If so, are these differences changing over time?

6. Health

6.1   The Population Census is quite limited in the collection of data directly relevant to health, although issues related to the work force discussed in Section 7 and 8 are particularly relevant to the planning and provision of health services. There has always been a significant proportion of people who have health-related qualifications but choose to leave the labour force or to work in other occupations/industries. A longitudinal data set would assist in understanding some of the factors affecting this. National shortages of people with health work force qualifications, skills and experience are sometimes acute at regional level, and an information base that supported both a fine level of geography and a longitudinal perspective on education/training/work patterns for health related occupations would provide valuable data.

6.2   The Department of Health and Ageing has several specific population groups to which it gives particular policy attention, including children and youth, the ageing, the disabled, those with Aboriginal/Torres Strait Islander origins, and those from different family types.

6.3   In 2006 Census it is proposed that there be the important addition of questions on the need for assistance with self care, mobility and communication, and having a health condition, disability etc which is the reason for this need. While this will not provide specific information related to the diagnosis or cause of the limitation/disability, it will provide data on the size and characteristics of the groups in the population who have short term and long term health conditions, disability, and age. If these questions are included in future censuses, it will be valuable to understand factors which affect changes in these measures. As health status tends to deteriorate with age, the measures would be expected to reflect that - but it will be important to identify the characteristics of those who previously reported needing assistance and having conditions, but subsequently do not; and persons who begin to report needing assistance/having conditions. For example, information about changes in the living arrangements, employment status, income and place of residence of persons in these groups would give policy and program relevance to these measures.

6.4   Another subgroup for which there would be policy interest in analysing census data sets longitudinally is those persons enumerated in health and aged care institutions. While the census records persons living at the same place one and five years previously, there is only limited information on the characteristics of those long term residents of institutional care, as they move through progressively more intensive types of care.

6.5   A third health related use of census data is the establishment of high quality estimates of population subgroups to support reliable analysis of disease - specific incidence and death rates over time. The starkest example of this is the inferior health status of Aboriginal and Torres Strait Islander populations. To be able to better measure and evaluate the characteristics of those persons who identify as Indigenous would advance progress on this issue.

6.6   In the absence of census data recording information about specific impairments, the Census would not support epidemiological breakthroughs of the type made by H.O. Lancaster in establishing the link between rubella and deafness, corresponding to a cohort born around 1899 when there had been a known epidemic. But by relating census data to cause of death data, and potentially to specific disease register data, similar advances could be expected in understanding the aetiology of some serious diseases, and patterns in the characteristics of people who suffer and die from them. There has been significant work done in this area in other countries including UK and NZ (REFERENCES [3] - [8]).

6.7   Significant progress has been made in epidemiological analysis and research in Australia, using the extensive (but imperfect) health data sets available. These include data on causes of death compiled by ABS as a by-product of the registration of a death with the State/Territory registrations; hospital morbidity data compiled as a by-product of hospital inpatient treatment; and the National Health Surveys and other health-related-household surveys undertaken by the Australian Bureau of Statistics.

6.8   Demographic and epidemiological analysis would be significantly enhanced by relating data from the Population Census to data on Causes of Death. These are:

(1) Using the self-identification of Indigenous status in the census to improve the quality of Indigenous mortality (and cause of death) data. (At present, Indigenous status may be recorded by hospital or funeral director staff, and is recognised to be understated. The census records Indigenous Status, as perceived by the individual or a member of his/her household and would therefore be more reliable.)

(2) Improved Indigenous deaths data would result in higher quality estimates of the Indigenous population.

(3) Using data on employment status, occupation, industry and place of residence derived from successive population censuses to develop a longitudinal data set which would be related to causes of death, which could then demonstrate the impacts of working life and geographical factors on mortality. Many deaths occur after retirement from the active labour force, so occupational data are not considered very reliable. Similarly, by the time of death a substantial proportion of Australians have moved to a new location from where they have lived much of their life, limiting the usefulness of geographical analysis which

might lead to strategies and policies which could support reduction or avoidance of behaviours that lead to health risk.

6.9   The Australian Institute of Health and Welfare has commented on this issue, as follows:

"The five-yearly National Census of Population and Housing is a very rich source of data that could be linked to other data to yield very important longitudinal epidemiological information. For example, linking census data to mortality data for the Indigenous population will yield a much better estimate of Indigenous mortality than relying on imperfect Indigenous data from death registration and population estimates. Also, linking Indigenous census records between two censuses would throw much light on the dynamics of Indigenous identification at censuses. Submissions have been made by various organisations to the ABS to retain census records for such and other epidemiology purposes. The potential of using linked data for statistical and policy research is immense and the NHIMG has developed guidelines for data linkage and for the use of linked data." (Australia's Health 2004; AIHW, forthcoming)

7. Education and Training

7.1   The Population Census contains valuable information on present attendance at an educational institution, highest level of school completed, and highest qualification (level and field of study). The census also collects data on a range of variables relating to the person, which are or may be relevant to educational outcomes, including sex, Aboriginal/Torres Strait Islander origin, country of birth of person and parents, need for assistance / health condition / disability, income, whether working, hours worked. Additionally, data items related to other members of the household, including Aboriginal/Torres Strait Islander origin, country of birth, labour force status/occupation and income of parent(s), and the number of persons in the household, are relevant to analysis of educational outcomes.

7.2   By creating a longitudinal data set derived from 2006 and subsequent Population Censuses, it would become possible to analyse educational outcome data (in terms of highest level of school, and post school qualifications) for persons attending school in 2006, and to relate those outcomes to their social characteristics at the time of schooling.

7.3   An important education policy issue is the improvement in education opportunities for children in rural areas. A longitudinal data set of this nature would provide the basis to better inform on present and changing pathways and outcomes for rural children through school to post-secondary vocational education and training and/or higher education. Some limited use is made of address five years ago, as an inadequate proxy for evaluating movement of school and post-school students.

7.4   The states and territories have now largely standardised assessment of literacy and numeracy at specified levels of education. If those data sets were to be brought into ABS and combined with Population Census data for those children, and related family data, that would provide a very valuable data set from which relationships between social factors and attainment could be analysed.

7.5   One limitation of cross-sectional data sets is the lack of information on flows (apart from what can be obtained by recall). An important aspect of education and training policy is evaluation of the transition to the world of work. A longitudinal data set derived from successive Population Censuses would, in time, support analysis of flows through school, post school education and training, and into unemployment or work, including occupation, industry, hours worked and income. It would also provide reliable longitudinal data on people who leave the labour force.

7.6   There are some attempts to measure the transition between education and work. Graduate destination surveys and a longitudinal survey of Australia's youth are limited by both the quality of the data they receive and the modest response rates, and have a quite limited scope, typically looking only a year or so after completion of higher education or vocational qualifications. These limitations would be significantly removed by a census longitudinal data set.

8. Work

8.1   The Population Census includes information about employment status, occupation, industry, hours worked, whether looking for work and available for work, as well as a range of variables of interest to labour market analysts and useful for defining particular population groups.

8.2   The Department of Employment and Workplace Relations has defined several population groups for policy attention, including:

- Disabled
- Lone Parents
- Long Term Unemployment
- Migrants
- Indigenous
- Mature Age Workers
- Young People

8.3   Many of these groups are definable in terms of Population Census variables, and a longitudinal data base derived from successive censuses would provide a valuable analytical resource from which to derive patterns and characteristics. The geographic dimension and regional aspects of employment and unemployment are also an important factor in the DEWR's work.

8.4   The transition from education to work, already discussed above is of close policy relevance in the context of work.

8.5   Having a census based longitudinal data set would significantly enhance the analyses that will be drawn from other longitudinal data sets relating to youth, migrants, beneficiaries and the more general HILDA. If it were possible to further augment the census data set with summary information derived from the Centrelink payment data base, by either identified or statistical matching, the potential for policy related analysis

relevant to the defined population group of high interest would become even greater. A particular policy issue for DEWR and DFACS is understanding social and employment outcomes for people who are on long term unemployment and disability support programs, and understanding barriers to participation.

8.6   One population group not adequately definable in terms of census variables is the long term unemployed. (Being unemployed at successive censuses is not necessarily an indicator of continuing unemployment, given seasonal influences). However, a combined Census/Centrelink data set would be very valuable in analysing the longer term outcomes for those people who have been receiving unemployment benefits for a year or more, in terms of their social characteristics, participation in training, and education/employment history.

8.7   The policy issues increasingly being discussed around retention of older persons in the labour force, moving through part-time work etc could be well served by analysis of longitudinal data to complement information derived from ABS cross-sectional surveys.

8.8   DEWR devotes substantial effort to occupation - related analyses, at reasonably detailed geographic level, to assist in evaluation of regional labour market policy. Census data are used to support this, and also to assess levels of skilled people who have qualifications in a field where shortages are evident or will occur, but are working in other occupations/industries or not in the labour force. (Examples are nursing, teaching.) A longitudinal census data set would assist the understanding of the pathways along which such people move, and assessment of whether patterns of departure from such occupations and/or re-entry to them are changing over time.

9. Economic Resources

9.1   In 2006, each person will be asked to report their total income, including wages/salary, government benefits, pensions, allowances and other usual income, as a single data item.

9.2   While there may be minor data quality imperfections (e.g omission of lumpy payments such as interest/dividends, even though there is specific prompting to include them), this question will provide a valuable measure of available economic resources, related to age/sex, employment/educational qualifications/occupation/industry, geographical location including rural/regional dimensions, and related to particular population subgroup, including families which identify as dependent on unemployed people, Indigenous people, lone parents, migrants and people with disabilities.

9.3   The census question does not identify the source of income, although questions on employment status, age, disability and family structure information may give some indication of possible source.

9.4   A key policy dimension, which is not available from other ABS sources or administrative data sources, is to understand the pathways and transitions for low income people/families who are able to increase income and move out of significant social disadvantage, and to better understand their characteristics. A census based longitudinal

data set would enable analysis of the characteristics of people who change their position in the income distribution, over time. Analysis of such a data set would potentially support policy option analysis related to both family and community and taxation initiatives. Linking of census data to the more detailed household surveys of income, expenditure and housing would provide a valuable data set for this.

9.5   Longitudinal analysis of Centrelink data can provide valuable insight into the characteristics of people who remain as clients of Centrelink. However, understanding the factors which enable people dependent on pensions and benefits to move out of that dependency may be aided by linking to census data.

9.6   There is particular policy interest in the economic resources of people living in rural communities. This is discussed in Section 15, below.

10. Housing

10.1   The census, being dwelling unit based, associates each person with a dwelling and has address recorded, although names and addresses have not been kept in the past.

10.2   There would be some interest in the changing characteristics of residents of dwellings over time, particularly amongst social geographers. However, their data requirements would broadly be met by data at CD level, without requiring longitudinal data in respect of particular dwellings.

10.3   Data items related to the dwelling are:

- number of registered motor vehicles garaged on census night
- number of bedrooms
- ownership by residents (7 categories covering owned; rented; other occupancy)
- if rented, who from (8 categories, including government, relative, employer)
- payment for rent, mortgage, site fee
- internet access

No particular policy imperative related to these data items would indicate that a dwelling based linked data set is a high priority.

10.4   However, there would be some policy interest related generally to type of housing, cost of housing and general well-being, and the types of housing histories followed by people belonging to several of the population groups listed in paragraph 4.4, if a longitudinal data base were to be developed using successive censuses. Over time, this would be much more valuable than the limited analysis that can now be done about people who state that they had different addresses 1 or 5 years previously. Such a longitudinal data set would also provide information about changes to other data items (income, employment status, occupation industry, family size, marital status, educational attendance, disability, provision of care), some of which may be factors associated with a transition in housing.

10.5   From past census data quality studies, there is a known deficiency in the numbers

of households which report that they are living in public housing. Given the significance of policy analyses related to this, linkage with data sets held by public housing authorities would be worthwhile to improve the quality of this data item. The administrative data sets would have additional information on duration of public housing tenancy.

11. Crime and Justice

11.1   The census does not include any questions directly related to crime, whether as a victim or a perpetrator.

11.2   An abbreviated data set is captured for persons in prison on census night: age (date of birth), sex, country of birth, Indigenous status, marital status, employment status, educational attainment, previous location.

11.3   There is significant policy interest in the issue of reformation or recidivism, but data sources are limited. The annual Prisoner Census can provide data about the characteristics of persons in prison in the same jurisdiction.

11.4   A Population Census based longitudinal data set which included persons in prison would enable analysis of the characteristics of persons who have continued in prison, and those who have completed prisons terms but are again serving sentence 5, 10, 15... years later, possibly in a different jurisdiction, compared with former prisoners who have not been imprisoned again. An analysis of characteristics of people known to have been in prison previously would also be of policy interest, in terms of their integration to the community, further education, entry to the labour force, etc.

11.5   Similarly an analysis of the socio economic background of people currently in prison, derived from a longitudinal census file, may provide information on the education, income, family background and other factors that may be associated with crime.

11.6   Unless a large sample is used (or the full population) to construct a longitudinal data set, the value of this analysis would be limited by the number of persons in it who are currently, or were previously, enumerated in prison.

12. Culture and Leisure

12.1   There are no data items specifically related to Culture and Leisure in the Census.

12.2   However, there is considerable interest in the characteristics of people who work in related industries. ABS has produced publications based on people who identify as working in an occupation or industry relevant to Sport and Recreation, and Culture in the census (Cat nos 4148.0, 6273.0). There would be some interest in knowing more about the changing employment patterns and career paths of these people over time, particularly in understanding where people once employed in these fields move to. However, given extensive unpaid employment in culture and leisure (and, even where paid, it may be a second job), the census is limited as a data source, and other more effective means might be used to gather information to address this, if it were to be seen as a priority.

13. Population Groups

13.1   Of the general population groups of particular interest noted in paragraph 4.4, census data enable identification as belonging to

- unemployed people
- Indigenous people
- children
- migrants
- older people
- people with low income
- people with disabilities

13.2   For all of these groups, there is major policy interest in social disadvantage, and that includes change in various characteristics over time, most notably in relation to income, employment, disability status, family composition and housing. For most groups, education/training is relevant too. All of these would be well supported by a longitudinal data set showing changes in status at 5 yearly intervals.

13.3   For retirees, and lone parents, census data are not definitive, although combinations of data items support an approximation for retirees. In respect of lone parents, living arrangements, shared custody etc have become more complex, and analysis of a longitudinal data set containing information on parents and their children over time would be expected to significantly enhance present understanding of the different characteristics affecting the dynamics of family relationship.

13.4   Of the population groups, only crime victims are not able to be addressed through census data. ABS conducts periodic household surveys of crime victims. If data from these surveys were to be linked with census data, that could provide indicative information on geographic mobility, housing, education, employment and income outcomes of crime victims over time - although analysis of the impact of being a victim of crime would need more information on, for example physical and mental health than could be supported by a Census-related data set. Significant advances here are likely to need a specifically designed project, drawing on data from administrative records and crime surveys.

14. The Regional Perspective

14.1   There is a significant policy concern about factors affecting regions. Sometimes "regional" is synonymous with "rural", but many regional issues are equally applicable to parts of metropolitan areas and to other larger urban centres and their surrounding areas. (Specific rural issues are discussed in Section 15.)

14.2   In external consultations as part of the development of a Rural and Regional Information Development Plan, a high priority information need is for better and more comprehensive information about regional migration and factors affecting it. The socioeconomic characteristics, skill levels (in the context of regional shortages of skills)

and a better understanding of reasons for migration into and out of regions are issues on which better information is needed.

14.3   Each census is eagerly awaited and extensively used for regional analyses, as most other ABS data sets contain insufficient geographical detail or are based on sample sizes too small to support estimates of acceptable quality for geographical areas below State/Territory. While cross-sectional census analyses are useful, there are limitations to their capacity to inform on changes, and factors affecting the dynamics. A census longitudinal data set with sufficiently large sample size would offer opportunities to analyse the drivers of change in the composition of different regions, in terms of industry and occupation, socioeconomic status, population composition and change (age/sex/family structure, and overseas and internal migration - both where residents have come from, and where former residents have gone to - over a longer time span than the present census questions support).

14.4   To the extent that census data could also be linked to administrative data, especially health services usage data, they would allow analysis of the changing health status of people in different communities and their needs for different types of services. A link between census and Cause of Death data sets would provide a beginning, and would become more valuable as the years cumulate.

14.5   Linkage with education data would allow analysis of long term outcomes for people who leave rural and remote communities for education reasons, compared to those who stay.

14.6   Similarly, a linked data base would support analysis of entry to the labour force, and, over time, changes in patterns of labour force involvement and career path, related to regional origins. It would inform on the extent to which people who leave regional areas may return to regions or move to other similar regions as they make transitions through education and work into retirement.

14.7   When linkage of census data with administrative data sets is feasible, the capacity to relate regional origins and residence to such issues as long term dependency on unemployment benefits, or receipt of various types of income support, including aged, disabled, youth allowances and benefits would be strengthened, and could be expected to be heavily used.

15. Rural Communities

15.1   Traditionally agricultural policy drew heavily on economic data - production, trade and financing. Several years ago, that expanded to include environmental data, as the impacts and relationships between past and present practices and present and future potential became more strongly articulated. Now, the policy focus has broadened to include social factors. Policy makers now want to take account of economic and social impacts on farm households, and rural communities. As well as needing information on the characteristics of people working in specific agriculture industries now and planning for the future agricultural labour force, there is a growing need for data to support social impact assessments, in terms of both individuals and farm households, and rural

communities, for a range of rural and industry adjustment policies.

15.2   While small area population census data provide a beginning, they are limited to the extent that many now involved directly in agriculture no longer live on farms, and many farms now produce a diverse range of agricultural commodities and services. Additionally, income earned "off farm", whether through employment in other industries or in providing contract agriculture related services to other farmers, is now much more significant and needs to be taken into account in evaluating farm viability. A more integrated, broader data set would support more informed understanding of dependency on the particular agricultural industries for which industry adjustment policies are being developed, and the resilience of those directly involved and their regional communities to change.

15.3   If population census data and agricultural census data were to be combined, the resulting data set would support much better informed analysis of the social impacts of any of the commodity, and providing information about the farming family - age/sex, family size, education (attendance, qualifications) employment, personal income, disability status, housing costs all are important variables in understanding rural families and communities.

15.4   A combined data set drawn from both the Agricultural and Population Censuses could usefully be supplemented by specific surveys to add to knowledge of

- farm management and decision making
- sources and significance of off-farm income
- extent and effect of multiple job holding.

15.5   The potential addition of administrative by product data sets such as those drawn from provision of health and educational services, and receipt of government benefits, would further enhance an informed understanding of needs for and access to services in rural communities.

15.6   Some further exploration of practical issues is being done by the Population Census Branch and Industry Statistics Branch. This would be important, to more completely associate households with the farms which they operate or invest resources and draw income. Particular issues are the correct identification of people who have significant off-farm employment as well as a significant farm workload, and farm operators who do not live on the farm property.

16.   Conclusions

16.1   This paper has identified a range of topics for which the information base has the potential to be significantly enhanced if a longitudinal perspective could be analysed from a reliable data base.

16.2   There is particular policy focus on transition, and factors which affect outcomes, but the lack of longitudinal data is a limiting factor in the analyses that have been done. On many of these topics recognition of this has led to commissioning and development of

various longitudinal surveys. While they offer capacity to explore topics in much greater detail, they suffer from relatively small samples and from non-response and potential biases. These have particular implications for some of the groups of greatest interest to policy analysts, particularly marginalised and/or highly mobile groups. A longitudinal census would be limited in the depth of topics covered but have very high coverage of the populations of interest.

16.3   The census also provides a fine level geographic perspective on the topics included in it. This level of granularity is rarely available from other ABS data sources. It can be achieved from some administration by product data, although they often suffer from the fact that the data are not produced with statistics as the main purpose, and so data quality does not always receive high attention. Additionally, the geographic parameter available is often derived from postal address, which has some limitations.

16.4   The Population Census is the ideal data collection vehicle for relatively smaller population groups, for which there would be significant cost or insufficient sample survey collections. For several of the specific population groups discussed in this paper, this is relevant, particularly where details by age/sex, geographic dissections, etc are relevant to the analysis.

16.5   While the potential for linkage described in the various options is wide, substantial gains in information available would be made under more limited projects. There is a need to balance the privacy aspects of linkage against the potential public benefit from enhanced information bases. Testing will be used to gauge community attitudes to varying options of census linkage, as well as appropriate protocols for ensuring the ongoing confidentiality of the data.

16.6   There are significant costs associated with establishing a longitudinal census and supporting its use. As well as being contingent on community acceptance of the approach, a longitudinal census would be contingent on strong demand for such information and adequate funding for its implementation.

References.

[1] ABS 2003: Census of Population and Housing 2006 ABS views on Content and Procedure. Information Paper (Cat No 2007.0)

[2] ABS 2001: Measuring Well-being. Frameworks for Australian Social Statistics (Cat No 4160.0)

UK References

[3] L Hattersley, R Creeser Longitudinal Study 1971 - 1991 - History, organisation and quality of data. . OPCS (Series L S No 7) (ABS Library - F 304.6 LON)

[4] M Rosato, S Harding, E McVey, J Brown Research implications of improvements in access to the ONS Longitudinal Study ONS Population Trends Spring 1998.

[5] A Dale Developments in Census taking in the last 25 years. ONS National Statistics.

NZ References

[6] Blakely T. Socio-Economic Factors and Mortality Among 25-64 Year Olds. The New Zealand Census-Mortality Study. *[Doctorate]. University of Otago, 2001.*   [PDF 1,445 KB]

[7] Blakely T, Salmond C, Woodward A. Anonymous linkage of New Zealand mortality and census data. *Aust NZ J Public Health 2000;24:92-95.* [PDF 3,016 KB]

[8] Blakely T, Salmond C, Woodward A. Anonymous record linkage of 1991 census records and 1991-94 mortality records: NZCMS Technical Report No. 1 *Wellington: Department of Public Health, Wellington School of Medicine, University of Otago, 1999*   [PDF 467 KB]

Canada References

[9] Bernier, J, Nobrega, K: Overview of Record Linkage. Proceedings of Statistics Canada Symposium 99 Combining Data from Different Sources.

Attachment 1 - Abbreviations

ABS	Australian Bureau of Statistics
AIHW	Australian Institute of Health and Welfare
ATO	Australian Taxation Office
BSC	Business Statistics Centre
CD	Census Collectors District
DEWR	Department of Employment and Workplace Relations
DFACS	Department of Family and Community Services
DIMIA	Department of Immigration and Multicultural and Indigenous Affairs
HIC	Health Insurance Commission
HILDA	Survey of Household Income and Labour Dynamics in Australia
NHIMG	National Health Information Management Group
NSC	National Statistics Centre
NZCMS	New Zealand Census - Mortality Study
OPCS	Office of Population Censuses and Surveys (UK)
ONS	Office for National Statistics (UK)

Information Development Plans -- adopting 'Best Practice' in data collection and management

Background

1.   The role of the Australian Bureau of Statistics (ABS), as specified in the ABS mission statement, is to "...assist and encourage informed decision-making, research and discussion within governments and the community by providing a high quality, objective and responsive national statistical service".

2.   In understanding how best to perform this role, the ABS acknowledges that it is impossible for one organisation to meet the totality of needs for statistical information across the community. Indeed, the ABS believes the role of providing a high quality, objective and responsive national statistical service is best met by the ABS taking a leading role, in consultation with external agencies, to improve the quality and quantity of statistical output available from any source. If this function is performed well, not only is the range of statistics available for decision making increased significantly, but a better understanding by all stakeholders of the breadth and quality of information available leads to more informed use of that data.

3.   When considering how the ABS can best perform this task, the Statistician has identified management, coordination and promotion of statistical best practice as an area where the ABS should take a leading role. Functionally, this role is articulated as the National Statistical Service.

4.   One way to enhance the role of a National Statistical Service is to promote the preparation of an Information Development Plan (IDP) for the various fields of statistics. The ABS has embraced IDPs as a key to developing statistical leadership and coordination in various fields of statistics therefore ensuring that the ABS keeps abreast of the statistical needs of Australia while engaging data custodians and users.

What is an Information Development Plan?

5.   IDPs aim to establish an agreed understanding of Australia's statistical priorities and establish a shared responsibility between the ABS and major users and data custodians for collaborative work that will address identified statistical needs.

6.   An IDP is an agreement developed as a joint effort between key stakeholders, that defines the suite of information required to support policy in a particular field of statistics. Each IDP embodies three kinds of knowledge, and shared commitment, to statistical

development activity:
- demand for information - a picture of the statistics that would ideally support informed design and evaluation of policy, other decision-making, research and community discussion,
- supply of information (including raw data that might be used to create statistics) - a picture of the existing data pool that might satisfy the demand for information, and,
- agreed statistical development activity - identified through the comparison of demand and supply, which defines and prioritises.

7. Information Development Plans are not new with the first IDP prepared by the Australian Institute of Health and Welfare (AIHW) in 1995. Progression of IDPs in the ABS commenced in early 1997, in the National Centre for Aboriginal and Torres Strait Islander Statistics, followed later by other areas in both the Population Statistics Group and Economic Statistics Group.

8. Currently, there are four completed IDPs in the public domain. These cover 'disability and housing and community services in the ACT', 'national public health', 'M easuring Learning in Australia - Plan to Improve the Quality, Coverage and Use of Education and Training Statistics' and 'national community services'. Each IDP has been informed by feedback from a wide range of stakeholders with ABS' involvement varying between being one of many stakeholders to being the lead agency.

9. To ABS's knowledge, there are currently 28 IDPs being worked on that are at different levels of maturity.

The IDP process

10. The process for developing most IDPs is an iterative one involving extensive consultation and collaboration with key stakeholders. Broadly, the development of an IDP in most cases involves identifying and consulting with stakeholders, developing a draft document, distributing this for comments, incorporating comments, seeking stakeholder agreement, re-drafting and finalising the document. Following this process the signed-off document is distributed to stakeholders (and in some cases published), progress is monitored and the document reviewed in a few years.

11. Key tasks in the process involve:
a) Defining the field of statistics
b) Understanding the debate
c) Defining the desired information set
d) Assessing the existing data pool
e) Deciding what information would add value
f) Mutual agreement to implement the plan
g) Evaluating the IDP

12. More information about these steps is provided in Attachment 1.

13. A continuing challenge throughout this entire process is for the ABS is to ensure that communication with stakeholders is coordinated and effective.

Examples of IDPs in ABS

14.   Three IDPs are mentioned below as examples of where the ABS and other agencies has achieved solutions to outstanding data needs by working together through the IDP process. While these IDPs are examples of where the ABS has taken the role of the lead agency, other successful IDPs have been lead and driven by other agencies. Each of these IDPs have some overlap of subject matter and significant liaison across each field of statistics is required to ensure a complimentary and consistent approach was taken with overlapping issues in each IDP.

The Rural and Regional IDP

15.   The Rural and Regional IDP is in the final stages of development. Following the identification of key rural and regional researchers, policy makers and agencies the ABS undertook 5 months of user consultation, including face to face discussions and receival of written submissions. Representatives from 75 agencies were involved in the consultation with the majority of these from state and territory government agencies. The primary purpose of these discussions was to gather information about their specific information needs and, in particular, to identify those areas where regional research and policy making are currently being constrained by lack of suitable statistics.

16.   A wide range of issues were raised at the various consultation sessions. Current demand for rural and regional data in Australia has primarily been driven by a small number of distinct social, economic and environmental policy/research themes; namely rural and regional quality of life, rural and regional economic growth, and the sustainable management of rural and regional natural resources. The maintenance and development of community strength (sometimes referred to as social capital) in rural and regional areas has also emerged as a key concern for policy makers and researchers interested in regional issues, and has thereby also led to increased demand for rural and regional data.

17.   Recommendations made in the IDP will reflect the above broad themes and will seek to specifically address the various sub-components of these themes identified by users as information priorities. In the development of recommendations, issues such as the statistical feasibility of addressing specific needs and the associated resources required will also be considered.

18.   At this stage it appears likely that recommendations for statistical development made in the IDP will focus on the following issues:
  • the characteristics and performance of regional businesses, and by extension the broader regional economy
  • the social and economic well-being of regional populations (and sub-populations), including measures relating to:
    • service availability and access
    • income and expenditure
    • health
    • employment

- community participation and volunteerism
- education, training, skills development and availability
- access to and use of information communications technology (business and household)
- environmental programs and practices

19.   It is recognised that the ABS is unable to address all user needs and that other agencies may be better placed to drive statistical development in certain areas. Consequently, a key part of the development process for this IDP will involve negotiations with other major regional data providers to establish both ABS and agency commitment to leading statistical development in agreed information needs areas. This process will be informed by input from the RRSNC Advisory Group, consisting of key Commonwealth and state government stakeholders, which will provide advice on the possibility and level of their agency's involvement in furthering statistical research on recommended projects.

The Agriculture IDP

20.   The Agriculture IDP is currently being prepared for final release, and agreement has been reached with stakeholders on the content and priorities of statistical development projects. Approximately 25 key stakeholders including DAFF, ABARE, BRS, Department of the Environment and Heritage, and a number of data custodians were consulted regarding the need for and availability of data relating to the agriculture sector.

21.   Key stakeholders were consulted individually and their policies and data requirements documented. A draft document was prepared and circulated to all stakeholders. This process of distributing and updating based on comments was repeated and a final version prepared for circulation.

22.   Current policy issues and in particular, areas which are constrained due to lack of suitable data, are addressed in detail in the document. Data needs are prioritised into high, medium and low priority needs with agreed future strategies proposed for each need. The IDP confirmed a number of issues already suspected, in that users confirmed that timeliness, accuracy and access to data were major issues for agriculture statistics.

23.   As a result the ABS has focussed on further improving the timeliness of it's major agricultural survey data, and has spent considerable effort ensuring that a move to a new population frame for agriculture statistics based on the Australian Business Register will not impact adversely on the accuracy or timeliness of the data.

24.   Some of the other proposals contained in the IDP include:
- testing of a strategy to link the Census of Population and Housing with the Agricultural Census,
- a proposal to classify 2005-06 Agricultural Census data to meshblocks (subject to user funding) to allow fine level geographic data to be released from the census,
- the addition of a question on the Agricultural Survey form accounting for land use made of all agricultural land.

The Water IDP

25.   The Water IDP is nearing completion, although further consultation is still required. With water being an area of increasing community and government interest the need for more frequent and timely release of data is paramount. The preparation for this information plan began by setting the scope where it was agreed the IDP should encompass the statistical areas that are core to the ABS water accounts. In general this includes the stocks and flow (supply and use) of water in the Australian economy and, as there is increasing interest, the economics of water use (eg assessing the value from water use, water trading, water pricing and the value of water infrastructure).

26.   Over the past 12 months the ABS has been involved in a series of bilateral meetings with key water data users including the National Water Initiative, the National Competition Council, the Productivity Commission, Murray-Darling Basin Commission, ABARE, CSIRO and the National Land and Water Resources Audit, as well as various state agencies. Following these discussions an audit of agencies that supply or use data relating to water use, the associated financial statistics, and water quality statistics was undertaken. This stocktake revealed that at least 8 national (additional to the ABS) and 20 state agencies were custodians of data which fell within the scope of the Water IDP. These data sources were then assessed against their suitability (ie relevance, accuracy, timeliness, accessibility, interpretability and coherence) to meet the current ABS water accounting and data needs. This process identified two main deficiencies; timeliness (data are usually collected sporadically and on an ad hoc basis) and coherence (data are often not comparable between jurisdictions).

27.   At this point the ABS has taken an introspective look at data requirements for water in relation to the IDP. The ABS conducted a forum with all key stakeholders to discuss the data gaps as identified by the ABS and the proposals to address these gaps. These proposals included the expansion of current data collections, improved output in relation to timeliness and economics of water use, and the integration of ABS water accounts with other major water information sources. To achieve this the ABS is required to take a more active role in liaising with other relevant agencies and, in particular, to promote the adoption of standards and coordinate reporting methodologies.

28.   The members of this latest forum have been asked to report back to the steering committee in relation to these objectives.

29.   It is anticipated that, the soon to be established National Water Commission (a statutory authority due to provide recommendations at the end of 2004 to address a number of water related issues) will expand to identify more general gaps in water data and the organisations best placed to fill these gaps. ABS will keep abreast of this work to monitor developments in data availability and use. This information will feed into future versions on the Water IDP.

What next?

30.   Progress of identified issues and proposed solutions will be regularly monitored to ensure that statistical development activities are consistent with the proposals contained in the IDPs over the life of the document. Where necessary, statistical development projects proposed in the IDPs will be fine tuned to align with current issues, as policies and issues evolve. Continued consultation with stakeholders is essential so that the ABS (as the leading statistical agency) keeps abreast of policies and statistical issues.

31.   After approximately 2 to 3 years, the current round of IDPs will be reviewed and extensive consultation will be undertaken with key stakeholders. The objective of this next round of consultations will be to update policies and issues, identify current or emerging statistical data gaps and potential duplication of data collections, and identify those areas where stakeholders can continue to enhance the national statistical service by being a proactive participant in the statistical arena.

Attachment 1

a) Defining the field of statistics
- initial work involves defining the boundaries of the IDP. This involves the preparation of a statement describing the field of statistics to be covered by the IDP. Once this statement has been prepared, and the key topics within the field have been identified, it is necessary to reach agreement with the key stakeholders on this clear definition.

b) Understanding the debate
- detailed user consultation is undertaken to ensure a complete understanding of the issues relevant to the field of statistics. It is essential to the success of the IDP that, from this early stage, all key stakeholders embrace joint ownership of the process. In consultation with stakeholders it is necessary to:
  - identify long standing issues that have hindered policy making in this field in the past,
  - list and prioritise recurring public policy issues,
  - list significant issues affecting the field in recent years, and, identify areas of interest in the future (next 12-24 months).

c) Defining the desired information set
- this process involves identifying the key areas of concern. The information is organised into a summary format, grouping similar issues and identifying links between issues. Stakeholders are again consulted to gain formal agreement to the information model. Following this, a list of variables appropriate to meet the needs of clients are compiled, including the reasons for the inclusion.

d) Assessing the existing data pool
- this involves undertaking a data stocktake within the field of interest and performing a quality assessment of the existing data sources. This stocktake is done in collaboration between the ABS and major stakeholders, individuals and groups. Once the data stocktake has been compiled it is possible to prepare a preliminary supply map that provides an overview of the available data and their characteristics.

e) Deciding what information would add value
- involves comparing the existing data pool with the information set required to support decision making. As well as affirming current collections, this helps identify data gaps, overlaps and deficiencies in the existing information, and provides the information necessary to allocate priorities against these deficiencies.

f) Mutual agreement to implement the plan
- this process involves discussions between stakeholders to identify who, and how best, to address gaps, overlaps and deficiencies. Decide, in collaboration with all parties, who will be the driver and coordinate the IDP, agree on allocation of responsibilities for the various actions identified to implement the plan, the timeframes for each task, the shelf life of the IDP and the resourcing of this activity. Once this has been outlined final agreement and formal sign-off from all key parties can finalise the process. This agreement should

include implementation of the plan's components according to their priorities. Any consultation with legislative, ethics and privacy organisations should be finalised at this stage. Over the life span of the IDP, each of the partner agencies should progress their responsibilities as articulate in the IDP.

g) Evaluating the IDP
- the final process involves setting up an evaluation mechanism and timetable including an agreement with stakeholders to regularly review goals. Improvements to any future IDP processes for this field of statistics should also be documented to ensure that the review, anticipated to occur in about 3 years, is successful.

Statistical Training and Consultancy

1.   The ABS' National Statistical Training Institute was discussed at the last AGSF held in May 2004, and sparked some interest from members. Courses such as *Thinking Analytically, Problem Solving and Story Telling* (TAPAS) were seen as beneficial in the context of most agencies. The ABS is continuing to slowly grow the Institute, however, a fully fledged program which accessible by government agencies is not available in the short term.

2.   Following the interest shown in the National Statistical Training Institute, this paper provides information on the statistical consultancy service and training courses currently provided by the Australian Bureau of Statistics for government agencies. The ABS has a well established training program to deliver a set of introductory and intermediate statistical skills. Note there are fees for the provision of these services.

Statistical Training

3.   The ABS provides a range of statistical training courses for developing statistical skills needed to undertake statistical collection and analysis. The courses teach practical skills through formal presentations and interactive discussion, combined with participant involvement in individual and group exercises.

There are two streams of courses. First, for people managing or conducting surveys, training courses are:

   • Basic Survey Design
This course provides a broad overview of all facets of survey development. It covers developing survey objectives, selecting collection methodologies, developing questionnaires, processing data and reporting results.

   • Managing Statistical Consultants
This course provides participants with the skills to effectively select and manage a consultant hired to conduct a statistical collection, statistical research or other data management related task.

   • Introduction to Data Management
This course covers common tools, methods, resources and procedures used in handling statistical data.

4.   Second, for people using statistics, training is offered in how to select and correctly use statistical information. These courses are:

• Turning Data into Information
This course aims to develop skills in interpreting statistical data and providing meaning written reports.

• Basic Statistical Analysis
This course develops skills in basic statistical and graphical data analysis techniques.

• Quality Informed Decisions
This course introduces the use of a data quality framework to evaluate the appropriateness of particular statistical collections to make informed decisions.

5.   The courses provide an introduction to a wide range of issues confronting users and producers of statistics. They are suited to people who want to develop a working knowledge of using statistics and data, but have limited practical or theoretical experience.

Statistical Consultancy

6.   The Statistical Consultancy area of the ABS provides agencies with methodological advice and assistance to help them in informed decision making. This assistance is generally in one of the following broad areas:

• Survey and Sample Design
Assistance can be provided to meet an agency's information requirements through survey design, i.e. advising on sample selection methods, survey weighting or questionnaire design.

• Data Analysis and Modelling
Statistical Consultancy can apply complex statistical techniques (i.e. testing hypotheses and statistical significance, measures of accuracy etc) to an agencies data.

• Reviews and Tender Evaluations
Existing or proposed surveys can be reviewed, taking into account the level of statistical rigour and the appropriateness of methodologies. Assistance in the evaluation of the statistical and analytical content of tenders, or in the preparation of tender specifications can also be provided.

• Data Management
Statistical Consultancy can provide assistance with the organisation of an agency's data holdings into a statistical framework.

7.   The NSS web site contains information about statistical courses delivered by government agencies. Any statistical programs can be added via the contacts page of the web site. More information on statistical training and consultancy services provided by the ABS is available from the web site, or by contacting Mark Lound on 02 6252 5907 or email mark.lound@abs.gov.au.