



# TECHNICAL NOTES

ON

## POPULATION STATISTICS AND INFORMATION SYSTEMS

December 1996

### *In this Issue:*

- ◆ **Page 1:**  
ISSUES IN THE DEVELOPMENT OF  
SPATIALLY REFERENCED POPULATION  
DATABASES  
(Uwe Deichmann)
- ◆ **Page 49:**  
USE OF ADMINISTRATIVE RECORDS IN  
POPULATION CENSUSES AND IN OTHER  
DEMOGRAPHIC AND SOCIAL STATISTICS  
(Sirageldin Suliman)
- ◆ **Page 16:**  
SAMPLING TOPICS FOR DISABILITY  
SURVEYS (Anthony G. Turner)
- ◆ **Page 55:**  
THREE CHALLENGES FOR EFFECTIVE  
UTILIZATION OF THE 2000 ROUND CENSUS  
DATA  
(Griffith Feeney)
- ◆ **Page 29:**  
HOW TO IMPROVE CENSUS INFORMATION  
PROCESSING (Carlos Ellis)
- ◆ **Page 64:**  
A REVIEW OF THE USE OF CENSUS TO  
INVESTIGATE DEMOGRAPHIC TOPICS IN  
ASIA AND THE PACIFIC COUNTRIES  
(Iqbal Alam)
- ◆ **Page 36:**  
PROPOSED CHANGES IN THE  
INTERNATIONAL RECOMMENDATIONS ON  
POPULATION AND HOUSING CENSUSES  
(Sam Suharto)

**Technical Notes**, issued by the Statistics Division of the Department for Economic and Social Information and Policy Analysis (DESIPA), is a collection of materials on population statistics and information systems prepared by staff of the Division. These notes present a summary of research by one guest author and the staff of the Statistics Division, primarily the UNFPA-TSS Specialists, in their individual capacities, and a review of recent developments in demographic statistics, data processing and other related fields. The views expressed in these Notes are, therefore, those of the individual authors and do not necessarily represent the views of the United Nations.

Readers are encouraged to provide feedback so that the Notes can be continuously improved. Please send your comments and suggestions to the Director of the Statistics Division, United Nations, Room DC2-1420, Two United Nations Plaza, New York, NY 10017 USA. Tel: (212)963-4996; Fax: (212)963-1940.

## THREE CHALLENGES FOR EFFECTIVE UTILIZATION OF THE 2000 ROUND CENSUS DATA<sup>13</sup>

Griffith Feeney  
Senior Fellow, East-West Center, Honolulu, Hawaii

*Note: This paper by Griffith Feeney was presented at the 17th Population Census Conference, held in Jakarta, Indonesia, 4-7 November, 1996. The Conference was attended by more than 40 national census directors, officials and experts from America, Asia, and the Pacific and from some international organizations. The theme of the Conference was "Information Technology and the Census of Year 2000." The content of the paper is quite relevant to developing countries which are planning and preparing for the 2000 round of censuses. Accordingly, it is being reproduced, unedited.  
(Sam Suharto).*

The three challenges mentioned in the title of this presentation are

- Capture everything digitally
- Exploit the unique comparative advantage of census data
- Integrate utilization of population census data with data from other sources

I will discuss them each in turn.

### ● **The First Challenge: Capture Everything Digitally**

At our last meeting in New Delhi, Sam Suharto presented a paper in which he suggested that computers should be used in all aspects of the census operation. I want to extend his point by suggesting that you should create a complete digital record of

every aspect of your 2000 round census operation, including, for example,

- planning documents
- questionnaires
- enumerator training materials
- codebooks and coding manuals
- all publications

It should be emphasized that I am speaking here only of management of the census information resource, not about dissemination to users. What materials you decide to disseminate in digital form, and in what media, are important but independent issues. With respect to dissemination, the only point here is that you will effectively foreclose the option of digital dissemination if you don't capture the census information in digital form as you produce it.

I would also like to caution against the idea that some of you may not need all this information in digital form. Plan for the future, not for the present! We are in a period of extremely rapid social changes driven by extremely rapid developments in information technology. Most of you are still three years away from taking a census, and the main results won't be available for a year or two after the census is taken. They will be in use until the results of the 2010 round census are available, roughly 15 years from now.

At the current rate of change in information technology, five years is a small eternity and more than fifteen years is an unimaginably large eternity. You and your users may not be ready for digital now, but you would be rash indeed to suppose that you

---

<sup>13</sup> Paper presented at the 17<sup>th</sup> Population Census Conference, Jakarta, Indonesia, 4-7 November, 1996. Email: gfeeney@hawaii.edu.

and they will not be ready during the next decade.

I would add that the least developed countries probably have the most to gain here. Preserving paper documents in an organized manner is much more difficult and costly than preserving digital documents in an organized manner. Thus you are more likely to preserve and exploit the census information resource if you do so digitally.

### **Digital Versus Print Documents**

The importance of capturing everything digitally may be driven home by enumerating some of the ways in which digital and print documents differ. Most of you know most of these things, but as Samuel Johnson said, "Men need more often to be reminded than informed."

To make the points concrete, let me call your attention to the print report I'm holding in my right hand, the BPS national report on the 1995 intercensal survey, and the computer diskette I'm holding in my left hand.

*Storability (compressibility)* This print report I'm holding is storable on any bookshelf, and it's not all that big. I've been bringing such reports back from my travels for 24 years now, however, and I've got a lot of them. A few years ago I came back from China with all the 1990 provincial census volumes I could buy, and piled one on top of the other they make a stack taller than I am. And I could only find reports for half the provinces. My office is full, the cabinets outside my office are full, and our library is full.

Digital documents take negligible space, and technological developments are constantly reducing the space required for any given volume of data. Moreover, the typical file of census data, whether consisting of individual records or of tabulations, may be reduced in size by about 80 percent using compression algorithms. Print documents don't compress, as we are all reminded each time we pack our suitcases after one of these meetings.

*Transportability* This BPS report is transportable. I can carry it back to Honolulu with me, or mail it. If I have more than a few, however, --the 27 similar volumes for the 27 provinces, for

example-- this transportation is either onerous or expensive or both. There is the cost of packing, a non-negligible task for 28 volumes this size, as well as the cost of mailing.

The same thing in digital form would be perhaps one thousandth of the size and weight, and with access to the Internet (fully available in Indonesia) we can send it just about anywhere in the world we would want to send it nearly instantaneously and for virtually no cost. If Internet connectivity is not available, we can still ship information in digital media for a small fraction of the cost of print media. The data on all 28 volumes of this report series, for example, would fit on one large capacity floppy disk that will fit in a shirt pocket.

*Reproducibility* This print report is reproducible. All we have to do is send some staff to the copy machine (which costs as much as a computer, if it is a small one, and a good deal more than a computer if it is the heavy duty variety), have them work away at it for an hour or so be prepared to pay 10-15 cents per page. The same information on this diskette here could be copied by any of us here in a few minutes on our computer.

*Searchability* You can search through this printed volume visually by flipping through the pages, and your search may be facilitated by using the list of tables at the front. And quite frankly, for casual browsing of small quantities of information, print documents are still greatly superior to digital documents. Display and battery technology are the two least impressive areas of development in information technology.

The appeal of the print document diminishes, however, when the number of them increases, as any of you who have had occasion to consult the complete set of publications of (say) the 1981 Indian census will gladly attest. They take up perhaps 20 feet of space on library bookshelves.

With digital documents you can search a single report in a few seconds and hundreds of volumes in a few minutes. If you had a suitably compiled list of tables, you could search all census publications ever produced in human history in the same time it would take you to briefly peruse this one printed report.

**Computability** I've been working with population census data for over two decades, and I rarely use any of the published numbers directly. Nearly every use involves using published as input to some calculation, often followed by plotting. The data in print documents is of no use to me until I get it back into the digital form in which, of course, it originated.

For very small volumes of data this is merely tedious. Even for extremely modest volumes of information, however, data on educational attainment for the 75 or so changwats of Thailand, for example, the cost in time and effort prevents much work from being done that ought to be done. In digital form, even extensive data for the nearly 3,000 country level units in China is readily usable on our desktop computers. In print form only, the expense of data entry (months of work, tens of thousands of dollars) reduces its value toward zero.

**Formatability** Much of the practical work of preparing census publications consists of formatting tables of many different sizes and shapes to fit on a single, fixed page size. We may not be used to thinking of fixed page size as a disadvantage of the print medium, but when we compare it to the flexibility of digital media we realize that it is a very severe restriction indeed.

The programs we use to view digital documents can "scroll" both horizontally and vertically, so that we may view a "page" of any size on a computer screen of any size. Digital publication can dispense entirely with the constraint of fixed page size. Indeed, the very necessity of a concept of "page" disappears in digital context. Dispensing with the "page" unit and working with the "table" unit facilitates the work of users working digitally, as well as producers, because data is presented in natural, "logical" units rather than artificial page units.

To be sure, variable table size in digital documents is an issue when we want "hard copy." The correct approach to this problem, however, is general purpose formatting software that formats tables of arbitrary size onto pages of arbitrary size, with full functionality for reduplicating titles and row and column labels on every page, numbering pages, and so on.

Systematic use of such general purpose formatting software would save tens of thousands of hours of programmer time in national statistical offices around the world. The innovation required here is a not a matter of software engineering. The functionality is already embedded in dozens of existing applications. What is required is innovation in management, organization, and planning on the part of data producers. Such software would be run by users and could be used for any data whatever. It could also be run by producers to produce print publications (though experience suggests that some "tweaking" will usually be required to get print documents into final form).

Fonts and font sizes are another aspect of formatability. For print documents, these are fixed at publication. For digital documents in suitable format, fonts and font sizes-- even colors!-- may be changed at will by the reader of the document. It is not necessary for digital publication to ape print publication's attention to the details of how documents look. Indeed, forcing a particular look on users is likely to make it harder for them to use the data.

**Printability** Well here's one area print documents have an advantage. They're already printed. It's not very difficult to print a digital document, of course, with some qualification for complexities of digital and print format. Note however that it is several orders of magnitude simpler and less expensive to get a digital document into print form than it is to get a print document into digital form.

**Security** The comparison between print and digital documents here is more complex. Digital documents *can* be as secure as print documents, but they won't be unless you understand the nature of the medium and take suitable security measures. There is-- let us not just admit it, but shout it out loud for attention (I'm not really going to shout here, this not being the Asian way)--a very real risk that digital documents will be a great deal less secure than print documents. Some reflection on facts familiar to us all will explain why.

Three kinds of security considerations are relevant here. First, data *loss*, by which I mean complete loss of the information resource. There is usually

negligible risk of this for print documents, because large numbers of copies are produced (250 is a "large" number in this case), and because there is considerable dispersion in space. We may lose some copies, perhaps even many copies, but we probably won't lose them all.

The situation for digital documents is very different. Because digital documents are so easily reproducible, copies are rarely stockpiled in the manner of print documents. Often only one copy (not counting backups) will exist. Backup copies are absolutely essential to protect against the risk of loss, and this means not only having backups, but insuring that they are current and that you can find them if you need them. This requires either an automatic backup system or a small dose of knowledge and a large dose of discipline on the part of your staff. In the absence of an automatic backup system, *every* one of your staff who uses a computer should have initial and periodic refresher training in backup basics.

The second security consideration is *corruption* of digital documents. A document is corrupted if it has been changed in some way that it should not have been changed, whether by careless error or malicious intent. The real danger of corruption is not that it happens, since we will always (!) be able to recover an uncorrupted copy from backup. The real danger of corruption is that we will not know that it has occurred.

It follows that we need a way of knowing whether or not a document has been corrupted. Programs that compare two files and report whether or not they are identical (such as the Unix *diff* or facilities provided by many PC utility suites) may be used in conjunction with backups. Validation may also be effected by the computation of *message digest functions*, such as Ronald Rivest's MD5, which assign codes to documents that will change (with extremely high probability) if the document changes.

Corruption is rarely an issue for print documents, because it is difficult to alter the contents of even one printed document, much less all extant copies, without the corruption being detected. Digital documents are nearly effortless to alter (one of their beauties, in other contexts) and usually exist in small

numbers.

Various measures are available to protect the integrity of digital documents.

- Have backup copies available for restoration
- Use operating system protections (e.g., make files "read only")
- Secure systems against computer viruses
- Have means of verifying integrity of digital documents
- Provide file access via secure server or intranet
- Use write-once media such as CD-ROMs

A related risk, less severe than corruption, but pernicious nonetheless, is multiple digital copies of a document that are supposed to be the same but are not. For the "stand alone" worker, the only protective measure is a bit of knowledge and a lot of self-discipline in what might be called "personal system administration." Put this on your list of training needs if such training is not happening already. For networked workers, maintaining a single (but of course backed up) authoritative copy on a file server, with a single responsible person able to make changes, is an very effective solution.

The third security risk is *escape*, by which I mean digital documents finding their way to places they do not belong. Sensitive print documents generally exist in few copies and are physically secured. Digital documents on non-networked computers may be physically secured in the same way, and may often be further secured with password protection.

Digital documents on networked computers require a very different (*digital!*) securing, about which we can say no more here than that there is a great deal to know and that it is absolutely vital to the integrity of your operations that your organizations have computer security expertise.

### How To (*very* brief)

Setting a goal is one thing, achieving it another. What steps can you take to capture the census information resource in digital form? We can't

descend into details, but here are a few overview points.

The simplest approach to capturing published reports digitally is to "print to file" at the same time you "print to paper." Files may be organized into directories and subdirectories, with a "readme" file in each directory. Close attention to the structure of file names is essential. Each published volume may take the form of a directory, for example, containing subdirectories for front matter, text, tables, and end matter, with the tables subdirectory further subdivided into subdirectories for tables on different subjects, and so on. This low (software) tech approach is very effective and can be done by every national statistical office right now. Achieving digital publication doesn't require new computer systems, it just (!) requires management, planning, organization, and training.

An elaboration of this approach that greatly improves the user interface with modest marginal investment of time and effort is to replace the readme files with HTML documents with links to each file (or parts of a file, if a single file contains many tables) and to higher and lower level directories. Users may then browse the digital publication using any web browser. The browser will display files with a .txt extension direct as text files. This method allows seamless inclusion of graphics as well as text-- without using any dedicated, proprietary and expensive electronic publication software. Non-text files, such as WordPerfect or Excel files, may be incorporated as well, though they will be readable only if users have the corresponding applications or a file viewer/conversion utility and configure their browsers to call the requisite "helper" applications.

It should perhaps be emphasized that using this method does not imply putting publications on the World Wide Web (www), though it makes mounting on a web site effortless if you decide to do so. It merely utilizes the HTML standard of hypertext links and the widespread free availability of web browsers to provide a convenient interface to digital documents for stand alone computers.

That's the low end for digital publication. What of

the high end? This is almost certainly a migration of publication processes to SGML, the International Standards Organization's Standard Generalized Markup Language. SGML is too large a topic to introduce even briefly here, but three comments may be useful.

- SGML is an internationally recognized, open standard (ISO 8879)
- SGML provides a highly disciplined way of structuring documents that distinguishes document *content*, *structure* and *format*
- SGML is familiar to many of you in the form of HTML, Hypertext Markup Language, which is a particular instance of SGML.

The explosive growth of the world wide web over the past few years shows the tremendous power of such open standards, voluntarily adhered to. For more information on SGML visit <http://www.sgmlopen.org>, which contains useful (and regularly updated) information, including links to other relevant sites.

### Think of Paper in a New Way

Many years ago I lived for several months in the old Ford Foundation Guest House in Kebayoran in suburban Jakarta. One of the other guests at the time was a Japanese professor of Economics, and he and I spent many hours discussing many topics over dinner during leisurely evenings. One evening the talk turned to Japanese politics, he made an acute observation. Japanese politicians do not like to change the law, he said, they prefer to *understand the law in a new way*.

We need to think of paper in a new way. Specifically, we should learn to think of paper

- as a computer interface
- as a way of looking at digital data
- as a convenient alternative to the computer monitor
- as something that will be with us for decades to come, but also

- as a medium that is dead already as a fundamental repository for new information

The qualifier "new" in the last point is essential. "Legacy" paper documents will be a fundamental repository for many decades to come.

How can a national statistical office lose 80 percent of the value of its data, if for some perverse reason it wanted to? Easy. Just print the publications and throw away the files.

## ● The Second Challenge: Exploit the Unique Comparative Advantage of Census Data

Population censuses are under attack today, both for their cost and for their relative lack of timeliness. Everywhere we turn these days we hear of nearly instantaneous communication, around the world, effortlessly crossing national boundaries. We learn that the secret of these wonders is digital technology, the movement and processing of "bits" (the smallest unit of information) rather than "atoms" (the smallest unit of matter).

What could be further from this model than a census enumeration, the fielding of an army of enumerators across an entire nation, reaching to the remotest and most inaccessible nooks and crannies of human habitation. Does all this movement of atoms still make sense as we enter the 21st century? Are population censuses dinosaurs of the information age, destined for extinction?

To address such criticisms we must think through the comparative advantage of the population census as a source of socially useful information. What is the comparative advantage of the population census? Quite simply that, as a complete enumeration, it can provide data for *every* subnational aggregate, and it can locate these aggregates in geographical space.

Having recognized the comparative advantage, we have to think through what is the specific *value* of this advantage. We have to review the traditional uses of census data and ask whether there are new opportunities for the use of census data that exploit

its comparative advantage and make it valuable despite its admitted weaknesses of cost and infrequency.

### Traditional Uses, New Opportunities

Uses of population census data may be divided into three broad categories. The first category is analysis of national, regional, and state/provincial level data. These uses do not in principle require a census, however. A large scale survey will generate comparably useful data at a small fraction of the cost of a census.

Why do I say "in principle?" Because saying "we don't need to put that in the census, we can get it from a survey" is one thing, actually getting the data from a survey quite another thing. The census has a visibility that sometimes makes it easier in practice for users to insure that the data they want really does get collected and tabulated.

The second category is local use of small area data. To the extent that all local areas want and need the same data, a census is necessary for this purpose. This is an important use of census data, but it is also problematic. Constituencies are local and diverse. Local areas often want to know *everything* about themselves and are dissatisfied with the content limitations of census data. This category may be extended to use of data for very small population groups, e.g., small ethnic minorities.

The third category of use is political representation. This is a very important use, obviously, and a census is required. Note however that political representation requires extremely limited data, often no more than numbers of persons. A population census that included only those items required for political representation purposes would be a radically impoverished census by comparison with the censuses taken around the world today.

Information technology has radically extended possible uses of population census data beyond the traditional models, but to realize these uses we must learn to exploit the power of digital technology for dissemination and utilization of data in the same way

we now exploit this power for producing data. Digital publication is the critical link here, and that is why I've emphasized the advantages of digital over print media.

Exploiting the unique comparative advantage of population census data requires producing modest to very large quantities of data that is highly focused toward particular uses. This is a radical shift from the generic, omnibus publication programs for census data that have characterized censuses for a century. Making the shift requires changing accustomed modes of thinking, managing and doing as well as mastering the technical details of producing useful digital output. Digital publication is critical, however, because the volume of information and the relatively limited number of users makes print publication hopelessly uneconomical.

Let me now suggest several new opportunities for utilization of the 2000 round of population census data.

#### Cities and Hinterlands

A major opportunity for the 2000 round population censuses is the provision of detailed data on the world's cities and their hinterlands. Uses include urban planning, environmental impact studies, and investment and marketing studies.

A census can provide vast quantities of data for individual cities-- but too often they do not. It is a terrible waste to produce large numbers of census tabulations for "urban" and "rural" areas only when the same tables could be produced for every city in the country.

Small area data *per se* does not provide useful information on cities. Cities are functional socioeconomic entities. Their boundaries often do not coincide with administrative areas. Moreover, the important area for many purposes is not the "city" but the "city-plus-its-hinterland." Understanding "city-plus-hinterland" requires knowing something about what geographers call *central place theory*.

Here as more generally, digital publication is the enabling condition. With the information technology now in place, most national statistical offices represented here can easily produce data for the ten, hundred, or, for that matter, the thousand largest cities and towns in the country. *Print* publication, warehousing, and distribution costs will usually be prohibitive. *Digital* publication is relatively inexpensive now, and costs will continue to decline rapidly for at least the next decade.

#### Population and Environment

I don't need to emphasize to you the importance of issues involving population and the environment. Studies of population and the environment mostly fall into two broad classes, global studies and locals studies.

The problem with global studies is that the world is too heterogeneous, with respect both to population and environment, to serve most studies well as a unit of analysis. Ignoring vast differences in (for example) physical environment, population density, and level of development obscures cause and effect. And understanding cause and effect is precisely what we must do to take effective action.

Local studies, on the other hand, are plagued by the problem of representativeness and the likelihood of adaptations spreading beyond local boundaries. How reliably can we generalize from one small Himalayan village (say) to the world? Communication and markets tie even small and remote local areas to broader regions. While environmental impacts are observed locally, understanding and correcting problems will usually involve looking beyond the local level.

What we need to better understand population and environment issues are *studies with high spatial resolution*, studies that look simultaneously at local, regional and national data by analyzing thousands or tens of thousands of small area units covering a large region.

Population censuses are not just an important source of population data for such studies, they are the

*only* source. No survey, no matter how large, can come anywhere close to providing the spatial resolution that every census provides. Indeed, the data potentially available for population and environment studies on the population side is vastly superior to what is available on the environment side.

The key word here, however, is *potentially*. We have not really begun to learn how to exploit the vast potential of using population census data to understand one of the most important public issues of this and the next century.

This is a tremendously important challenge for the 2000 round population censuses. It cannot even be contemplated with the traditional census procedure of dissemination of printed reports. The data involved is so voluminous that most of it should probably never be printed at all, much less in reports with multiple copies.

The necessity of articulating population census data with environmental data requires a new model of interaction between data producers and data users, a model in which they work together not just on the content of the census products, but on the details of the digital interface.

### **Example 1: Water Resources**

Let's look at two examples, water resources and global warming. First, water resources. For many countries, water may be the most critical resource of the 21st century. The problem, as with food, is not the total amount of water available, but its distribution in relation to population.

To study water resources in relation to population (or anything else) we need to use the appropriate spatial unit. The appropriate spatial unit is the *drainage basin*. Water moves according to the law of gravity, not according to administrative classifications. When was the last time you tabulated your population census data by drainage basin? Never, for most of us.

We in population census work have become much too fixed on administrative boundaries as the

appropriate spatial units for presenting census data. It is a mental habit we must break if we are to exploit the unique potential of census data, which *unlike any other source* can be tabulated for any spatial unit.

The practical way to realize this potential for most countries is not geocoding households, or even enumeration districts (though this makes sense for some countries). The practical approach is to produce data--digitally!-- for large numbers of small administrative units and approximate larger spatial aggregates by grouping the small administrative units together.

The resolution—the number of units—required is very modest by the standards of today's computing power, but very large by the standards of most traditional census presentation. Large countries should be thinking in terms of tens of thousands of small administrative units at a minimum.

### **Example 2: Global Warming**

How should society address the impact of a possible rise in sea level? Rather obviously, one of the first things we want to know is how many people would be directly affected. The population census contains the data, but we need to tabulate it by proximity to coastline and by elevation.

Tabulations of population by elevation of residence are distinctly uncommon in census work. Have you ever seen one? I haven't. We should see them following the 2000 round of censuses. Initially, at least, the practical approach will probably be to approximate purely topographical spatial units defined by elevation by aggregates of large numbers of small administrative units (or perhaps by census enumeration district).

### **The Need for Advance Planning**

I've said it more than once already, but it is so important that I want to say it again under its own section heading. Exploiting these new opportunities for the utilization of population census data requires advance planning involving producers and users, far

more than in the past.

The digital interface between producers and users is at present very undeveloped. To realize the potential created by information technology it needs to be developed much further and understood much better on both sides. The time to begin planning for this is *now*, while there's still time to make the necessary adjustments in data processing procedure.

### **The Third Challenge: Integrate Utilization of Multiple Sources**

Data producers learn to think in terms of what must be done to collect information of a particular type. We learn to think in terms of

- sectors (population, agriculture, industry, etc.)
- instruments (census, survey, register, etc.)
- operations, production, dissemination

Some data users learn over time to think in the same terms, and in the nature of the situation, these are probably the users you tend to interact with most.

In general, however, users do not care about "census data," or "survey data," or even "agricultural data." They want *information* relevant to their purposes and concerns. It is not so easy to deal with these users, precisely because they don't think in our terms. They just want to know about homelessness, or AIDS, or the condition of women. They don't care about sectors and instruments, and they want (and deserve) all available information, not just information from a particular source like the population census.

Where producer and user perspectives diverge, both prudence and duty recommend that we make every effort to overcome our natural attachment to the producer perspective and try to adopt the user perspective. When we do, we usually find that

addressing user information needs requires drawing on data from different sectors collected with different instruments. We are also likely to find that putting data generated by different parts of our organizations together is not so easy.

Effective utilization of census data requires

- attempting to adopt the perspective of key users
- which means focusing on *issues*
- which usually means pulling in "disparate" information sources
- and figuring out how to make this information work together to serve the users interests

Paradoxically, effective use of census data requires forgetting about "census data" as such, at least temporarily, and focusing on issues and users.

The examples given above suggest several non-census data sources that need to be integrated into census analysis.

- administrative boundary maps for geocoding and GIS
- environmental data (land cover, pollution, etc.)
- topographic data (coastline, rivers, elevation)

### **Conclusion**

- Population census data can make major contributions to understanding the critical issues of our time, but ...
- We need to learn to use the data in new ways
- Far more than in the past, we need to plan utilization far in advance
- The time to begin planning for the use of the 2000 round population census data is *now*, before the data are collected and processed.