Appendix B. A Guide to Census Geography and Applications

Introductory Note

This appendix contains information previously available separately as the 1991 Census publication, *Geography Guide Book* (Catalogue No. 92-310). It is intended as a guide for those who are not familiar with the use of Statistics Canada's data in a geographic context. For simplicity's sake, some of the more complex features are excluded or are treated very briefly. Statistics Canada staff in any regional reference centre would be pleased to answer questions applying to your situation.

The five case studies provided in this appendix give specific examples of how to tap into the geographic aspects of common types of data requests. While the specific names and research cases are fictitious, the types of situations depicted are very common. The cases are simplified and focus only on the geographic aspects of the situation. In most cases, many other research and business considerations are also important but are not explored.

How Statistics Canada Organizes Its Census Data Geographically

Questions such as: what's happening in this province? this market area? or this neighbourhood? require answers with geographic perspectives. The ability to produce data for specific geographic areas depends upon the way the data are collected, stored and retrieved. If the measurement is of something that occurs throughout Canada and surveys everyone, then very detailed geographic information can be produced. For this reason, the Census of Population and Housing releases the most detailed data for small geographic areas available of any data program produced by Statistics Canada. This appendix discusses the geography of the 1996 Census of Population and Housing.

The Beginning

The Census of Population and Housing (hereafter called the census) collects data on every person in Canada based on where he/she lives. The basic unit of measurement is the dwelling. In other words, the census begins by locating all places of residence in Canada. For each "private occupied" dwelling, there is one household which consists of one or more persons. (Special procedures have been put in place to deal with persons who do not live in a dwelling.) For subsequent tabulation purposes, each dwelling is given a series of geographic identifiers.

The first geographic identifier assigned is a direct result of the initial collection step. Since all dwellings are located by a physical search of an area assigned to a census representative, that area is given a specific geographic identification; it also has specific geographic boundaries to ensure no double counting or gaps. In rural settings, this "territory" can cover wide reaches of land but in urban settings, it is usually several city blocks. Because the Census Representative used to be called an enumerator, this area is called an ***enumeration area*** or EA for short. All dwellings in the enumeration area are given the enumeration area's unique geographic code (this assists in classifying geographic areas).

Until 1971, no further geographic breakdown was available. Now, within larger urban centres, a more detailed geographic identification systems is used that permits smaller geographic presentations of the data. It is called the ***block-face***, which usually refers to one side of a street between two consecutive intersections or major intersecting geographic features such as railway tracks. In effect, it is a small strip of houses. For large apartment buildings, separate block-face designations are assigned if the apartment building constitutes a complete EA. In fact, a very large apartment building can be subdivided into more than one EA, and each EA will have its own block-face.

Block-faces in large urban centres and enumeration areas in smaller urban centres and rural areas provide the finest geographic detail available for almost all geographic applications. In other words, when tabulating census data geographically, all households in the basic unit (block-face or EA depending on the area's location) must be either included or excluded. Case C, indicated later in this appendix, deals with the one limited exception.

For many applications, it is easier to regard the EA or the block-face as a single point rather than a small area or strip. Thus a ***representative point*** is identified that is suitably located within the EA or block-face. Examples of block-face representative points are shown in Figure 21.

The Geographic Hierarchies

The census organizes the dissemination of data based on hierarchies of geographic units. There are three different hierarchies because of differing origins of the geographic units. There is a national hierarchy, where each level of geographic unit covers the entire country, a metropolitan (urban) hierarchy where each level of geographic unit applies only for urban centres, and a postal code hierarchy for the convenience of users requiring census data by postal code geography. The geographic hierarchies include several levels, some of which nest completely within the next larger level and others that do not. By "nest", it is meant that adding all smaller units within a larger unit leaves no part uncovered nor does it result in any overlap. Some levels appear in more than one hierarchy. The hierarchies and their interrelationships are depicted in Figure 20.

Understanding the hierarchies and their interrelationships is important for accurate data retrieval from census databases. The hierarchies are implicit in the geographic codes used to access census data and it is usually necessary to use the codes from two or more levels in the hierarchy to ensure the exact geographic units of interest are obtained.

The National Hierarchy of Geography

The top level in the national hierarchy is Canada and the bottom level is the enumeration area (EA). The EA is defined to respect all higher levels in the hierarchy and is therefore often referred to as a "building block". In between these two levels, there are provinces and territories and many subprovincial levels which are described in more detail below.

While not depicted in Figure 20 as a separate level in the national hierarchy, the provinces and territories can be grouped together by "region", such as the Atlantic provinces or the Prairies. These "region" groupings are identified by the first digit in the two-digit geographic code for provinces and territories. (For further details, refer to the definition for ***Province/Territory***.)

The constitutional basis for the census originates from the requirement to apportion federal electoral representation based on population counts. Thus, one of the levels in the national hierarchy is the federal electoral district (FED), the federal Member of Parliament's riding. Enumeration areas are defined to respect the FED boundaries and the FEDs add together to form provinces and territories. (For further details, refer to the definition for ***Federal Electoral District***.)

Many provinces are already divided into official areas for regional and local government purposes. Most of us are familiar with terms such as counties, regional districts, regional municipalities, municipalities, townships and Indian reserves when referring to these subprovincial administrative areas.

Census division (CD) is the general term applied to areas established by provincial law which are intermediate geographic areas between the municipality and the province levels. Census divisions represent counties, regional districts, regional municipalities and other types of provincially legislated areas. In Newfoundland, Manitoba, Saskatchewan and Alberta, provincial law does not provide for these administrative geographic areas. Therefore, census divisions have been created by Statistics Canada in cooperation with these provinces for the dissemination of statistical data. In the Yukon Territory, the census division is equivalent to the entire Territory. (For further details, refer to the definition for ***Census Division***.)

Census subdivision (CSD) is the general term applying to municipalities (as determined by provincial legislation) or their equivalent (for example, Indian reserves, Indian settlements and unorganized territories). In Newfoundland, Nova Scotia and British Columbia, the term also describes geographic areas that have been created by Statistics Canada in cooperation with the provinces as equivalents for municipalities for the dissemination of statistical data. (For further details, refer to the definition for ***Census Subdivision***.)

According to the national hierarchy, census subdivisions add together to form census divisions. The CDs form provinces and territories. Two additional levels are defined in the national hierarchy to facilitate special data analysis. A special aggregation of census subdivisions called ***census consolidated subdivision*** (CCS) provides a level of geography between the CSD and CD which facilitates data analysis. In the rural context, the CCS is a grouping of smaller municipalities, usually contained within a larger municipality. For instance, a town located within a surrounding township will be grouped together with the township to form a CCS. In urban areas, CCSs are formed by contiguous groupings of CSDs. A principal user of the CCSs is the Census of Agriculture.

Agricultural data programs use subprovincial aggregations called ***census agricultural regions***, also known as crop districts in the Prairie provinces. Census agricultural regions are made up of groups of adjacent ***census divisions***. In Saskatchewan, census agricultural regions are made up of groups of adjacent ***census consolidated subdivisions***, but these groups do not necessarily respect ***census division*** boundaries. (For further details, refer to the definition for ***Census Agricultural Region***.)

There is another level in the national hierarchy that is used primarily for the dissemination of economic data. An ***economic region*** is a grouping of complete ***census divisions*** (with one exception in Ontario). Prince Edward Island and the two Territories each consist of one economic region. Economic regions are used to analyse regional economic activity. (For further details, refer to the definition for ***Economic Region***.)

A new level in the national hierarchy has been identified for the 1996 Census, but it does not nest with higher levels in the hierarchy. ***Designated places*** refer to areas created by provinces to provide services and to structure fiscal arrangements for submunicipal areas which are often within unorganized areas. The concept of a designated place generally applies to small communities for which there may be some level of legislation, but where the communities fall below the criteria established for municipal status, that is, they are “submunicipal” or unincorporated areas. Provincial governments require census data in order to administer grants and/or services to designated places. Prior to 1996, Statistics Canada facilitated the retrieval of census data by delineating these areas at the ***enumeration area*** level only. The increasing demand from provinces for population counts by designated places led to their recognition as a new dissemination geography for the 1996 Census.

Statistics Canada relies on provincial authorities to identify those areas to be defined as designated places and to provide adequate boundary descriptions or maps. As a result, the areas recognized as designated places may not represent **all** places having the same status within a province.

The Metropolitan Hierarchy of Geography

Most of Canada's vast land area is sparsely populated and, with each passing decade, a greater proportion of the total population is found in urban settings. In fact, more than 70% of Canada's population lives in urban centres with a population of 10,000 or greater. Based on certain rules with respect to population and density, all land is either urban or rural.

Urban-focused economies tend to expand beyond official municipal or even county boundaries in terms of shopping trips and commuter travel. As a result, Statistics Canada has created groupings of municipalities, or ***census subdivisions***, in order to encompass the area under the influence of a major urban centre. Specific guidelines are used to group municipalities that are closely interconnected due to people working in one municipality and living in another. The resulting geographic units are called ***census metropolitan areas*** (CMAs) for larger urban centres (100,000 or more in their urban core in the previous census) and ***census agglomerations*** (CAs) for smaller urban centres (with an urban core of at least 10,000 but less than 100,000 in the previous census). In the 1996 Census, there are 25 CMAs and 112 CAs. (For further details, refer to the definition for ***Census Metropolitan Area***.)

Beginning with the 1986 Census, the CMA/CA concept was further refined to accommodate the cases where the area of influence of one CMA or CA continues to expand to nearby CAs. When this happens, Statistics Canada identifies the individual components as ***primary census metropolitan areas*** (PCMA) and ***primary census agglomerations*** (PCA) which, together, form a larger ***consolidated census metropolitan area*** or ***consolidated census agglomeration***. Thus, in selected CMAs, there will be a PCMA and at least one PCA. For instance, the Edmonton CMA in Alberta is composed of the Edmonton PCMA, the Leduc PCA and the Spruce Grove PCA. Note that although Edmonton CMA is consolidated, we call it a CMA.

In CMAs and CAs, land is designated as being part of an ***urban core***, an ***urban fringe*** or a ***rural fringe***. Certain rules with respect to population and density are used to make the urban designations. Outside CMAs and CAs, land is also designated as ***urban area*** using the same rules. Area not designated urban is considered rural.

Users often need data for areas that are smaller than a municipality. As a result, Statistics Canada created ***census tracts*** (CTs) to equal neighbourhood-like areas of 2,500 to 8,000 people (preferably close to 4,000) within all CMAs and CAs that contain an urban core with a population of 50,000 or more in the previous census. The CT boundaries generally follow permanent physical features such as major streets and railway tracks and attempt to approximate cohesive socio-economic areas. One unique feature of CTs is that their boundaries are generally held constant from one census to the next, so that CTs are comparable over time. A subsequent census may split a CT, but normally it can be easily aggregated to equal earlier boundaries. This characteristic, however, means that CTs do not necessarily follow CSD or CD boundaries. This lack of nesting occurs most frequently when neighbouring municipalities adjust their boundaries between censuses. Only at the external outline of a CMA or a CA does a CT boundary have to follow that of a CSD or CD. In practice, however, there are few cases of CTs not nesting perfectly within CSDs and CDs.

CTs have had a long history of being the unit of choice for analysing neighbourhoods since data have been readily available for them (and for a long time, they were the only submunicipal data released in readily available form).

Postal Code System

The postal code system is a geographic system designed by Canada Post solely to facilitate the delivery of the mail. It is quite different from the geographic systems used by Statistics Canada.

The postal code system is hierarchical in that it builds from small units of location (often a block-face in urban centres) to a larger territory. The familiar six-character alphanumeric postal code is a reflection of that hierarchy. The first three characters refer to the forward sortation area (FSA) which, in urban areas, is about the size of four to six census tracts. The last three characters are the local delivery unit (LDU) which, in urban areas, is often equivalent to a block-face (normally one side of a city street between two consecutive intersections). There are many LDUs in each FSA. In areas with carrier delivery, groupings of LDUs form the carrier's delivery route, called a postal walk (PW), which is larger than an EA but smaller than a CT. Unlike census geography, which is "frozen" for five years, postal geography is constantly changing.

Most postal area boundaries and routes have very little correspondence with census boundaries even though they usually build up from the same base in urban areas (the block-face). In addition, the many delivery complexities, such as post office boxes, community mail boxes ("super-boxes"), heavy volume mail users and rural routes, make it difficult to always be able to fit postal geography into Statistics Canada's geography or vice versa. To assist postal geography users, Statistics Canada has created a linkage file called the Postal Code Conversion File (PCCF). The PCCF indicates in which EA (or EAs) each postal code (FSA-LDU) is located, where possible. It also shows the representative point or representative points for the postal code, based on the block-face or the EA information. The PCCF facilitates the analysis, with data coded by census geography, of information that is coded by postal codes. Users should be aware that postal geography and census geography do not match perfectly.

Other Geographic Systems

Many other ways of organizing Canada have a spatial context including telephone exchanges, soil zones, broadcast listening areas or "footprints", election polling stations, provincial ridings, municipal election wards, climatic zones, drainage basins, and so on. None of these necessarily fit easily into census geography. However, it is still possible to compile reasonable estimates of census data as long as lines can be drawn on a map. Case C explores these types of situations more fully.

A Note on Numbering and Naming

The main components of the national hierarchy (CSD, CD, provinces and territories) are reflected in the Standard Geographical Classification (SGC). This is a three-tier coding system that assigns a unique number to each area covered by the classification.

Additional numbering systems are used for EAs (linked to the FED and province or territory code), CTs (linked to the CA/CMA code), CMAs/CAs (including the province code, which is unique except for British Columbia, the Yukon Territory and the Northwest Territories, that begin with the digit 9). Manuals which list all the codes, and the location of the areas on maps, are available for reference purposes.

Statistics Canada follows certain naming conventions. All municipalities that are recognized governmental units have their officially designated names listed in the SGC. For geographic units designated for statistical purposes (EAs and CTs), numbers are used. Geographic groupings of a number of geographic units that are designated by Statistics Canada, such as CMAs and CCSs, take the name of the most prominent locale.

Key Considerations for Working With the Census Geography

As with any research endeavour, there are always caveats and pitfalls to be aware of when working with data and their geographic elements. The following are the most notable.

Boundary Revisions

Geographic boundaries for cities, townships, etc. are "frozen" by Statistics Canada in the year of the census (normally on January 1 of the census year). Therefore, any boundary changes implemented by a city or town after that date will not be reflected in the geographic tabulations of census data. For instance, if a municipality annexed a portion of an adjoining township on March 15, 1997, the official 1996 Census results for that municipality will reflect the previous boundaries (as will all reference maps from the census), even though the data were released later in 1997 or in 1998. A special calculation would be done by Statistics Canada to put the population data on the new boundaries (see Case E).

Statistics Canada redefines, at each census, the boundaries for geographic units that it is responsible for defining (such as EAs, CMAs, etc.). These alterations are made in order to improve collection efficiencies or because of growth and change in urban areas.

Changes Since 1991

The total number of CMAs for 1996 remains at 25, the same as in 1991. There are 112 CAs for 1996, down from 115 in 1991. Two new CAs were created, but three CAs from the 1991 Census were dropped because their urban cores dropped below the minimum 10,000 population requirement. Two CAs were converted into PCAs.

Since census metropolitan areas and census agglomerations are meant to enclose the extent of urban influence of a market area, irrespective of administrative boundaries, their outer boundaries will expand if the urban area expands. Thus, data on a specific area by its boundaries from one census may not be comparable to data for the area with the same name from the next census. Similarly, changes in administrative boundaries (such as annexations) between censuses will provide comparability problems. Statistics Canada does indicate in its published data that changes have occurred.

Random Rounding and Area Suppression

For all census products released, procedures are applied to prevent the possibility of associating statistical data with any identifiable individual; the data are randomly rounded and they are suppressed for certain geographic areas.

Random rounding is a method whereby all figures in a tabulation, including totals, are randomly rounded (either up or down) to a multiple of "5", and in some cases "10". This technique provides strong protection against direct, residual or negative disclosure, without adding significant error to census data. However, figures on population counts only are not rounded since they provide no information on the characteristics of these populations.

Area suppression results in the deletion of all characteristic data for geographic areas with populations below a specified size. Thus, areas with a population of less than 40 persons are suppressed. However, if the data are released at the postal code level or forward sortation area (FSA) level, those areas with populations below 100 persons are suppressed. If the data contain an income distribution, those areas with populations below 250 persons are suppressed. In all cases, suppressed data are included in the appropriate higher aggregate subtotals and totals. For more information, please see the *1996 Census Handbook* (Catalogue No. 92-352-XPE).

Day-time Versus Night-time Demographics

Another important consideration is the concept of where the activity occurs in relationship to where it is measured. For instance, if a retailer wants to know the market nature of customers in the Portage and Main District of Winnipeg (in the heart of Winnipeg's downtown), he or she would get a very incomplete picture using only the census data for the neighbourhood surrounding the site. By combining census residence data with census data tabulated by workplace location, the analyst can get a much more complete picture of the area of study. This phenomenon is sometimes referred to as "day-time demographics" versus "night-time demographics". Details on census workplace location are provided in the journey-to-work section of the dictionary.

Cottage and Tourist Areas

Another example of population counts reflecting permanent residents is in the cottage and tourist areas. In these areas, the census only includes permanent residents. Therefore, the demographic data for these areas reflect year-round residents.

Commonly Used (But Not Official) Names of Places

Statistics Canada follows officially established names and boundaries according to each province's and territory's Municipal Act or its equivalent. These names and boundaries often are not the same as those of many places that are treated as separate entities by historical custom or by organizations such as Canada Post. For instance, in Metropolitan Toronto, such areas as Don Mills, Willowdale and Mimico, no longer exist as official municipalities and, therefore, data for these places are not available as they would be for census subdivisions.

Similarly in many rural areas, a locality might not be an official municipality according to its province or territory. However, many people still use these places as postal addresses. Statistics Canada has traditionally published a listing of such places in rural areas with basic population counts calling them "unincorporated places". Such listings are not available as standard products for 1996. Additional demographic detail for these areas can be calculated by using EA data as described in some of the cases later in this appendix.

There are a number of other general guidelines to follow when using data divided into geographic areas:

1. Be careful not to double count if mixing different geographic units. For instance, if adding CMAs and CDs (counties) together from the same province, ensure that there is no overlap between the geographic coverage of the two. It is impossible to double-count when using the same type of geographic units (except when mixing PCMAs and PCAs with the full CMA or CA).

2. Beware that many geographic areas may have the same or a very similar name. For instance, Peterborough is the name of a city, a county, a census agglomeration and an urban area in Ontario. Whenever Statistics Canada products use a place name, they indicate what type of place it is, and in which province it is located.

3. Beware that some geographic areas cross the limits of other standard geographic units such as provincial borders. For example, Ottawa - Hull CMA crosses the provincial limits of Ontario and Quebec. Therefore, if you wanted to calculate the proportion of Ontario's population living within CMAs, you would not take the total for the Ottawa - Hull CMA into account, but only the Ontario part.

Two municipalities straddle provincial boundaries: Flin Flon and Lloydminster. However, these CSDs have a Standard Geographical Classification code for each provincial part. Consequently, the two parts must be aggregated for the total population.

4. When using data from different statistical programs of Statistics Canada, or from other data producers, ensure that the geographic definitions are consistent, even if the data are for the same time periods. For instance, the Labour Force Survey may use CMAs delineated from different censuses.

5. There are special situations where an EA will appear to have a relatively large population but only one household and one dwelling. These are "collective EAs" which usually apply to an institution such as a retirement home, chronic care facility or jail. Care should be taken to ensure that collective EAs are identified when using census data that include these areas.

6. Users should be aware of the geographic implications of obtaining data for non-standard areas. Authorized secondary distributors (see Box) provide a number of products and services based on tabulating randomly rounded EA data. Only Statistics Canada has access to the full range of census data by block-face and to unrounded, unsuppressed data. When dealing with non-standard areas that divide many urban EAs, but are about the same size as an EA or two, a special tabulation from Statistics Canada will yield more accurate data than a special tabulation from a secondary distributor, due to Statistics Canada's access to unrounded, unsuppressed data. The problem diminishes in relative severity as the number of enumeration areas aggregated increases.

The Role of Secondary Distributors

Statistics Canada has licensing agreements with a number of organizations to distribute 1996 Census data, including geographic information. This is an important component of the marketing and dissemination of 1996 Census data. In fact, only authorized distributors may redisseminate census data since all data (including boundary files) are purchased on an end-use basis, unless otherwise approved by Statistics Canada. A list of approved secondary distributors is available from your nearest Statistics Canada regional reference centre upon request.

Empty Places?

There are a number of EAs, block-faces and some CTs and CSDs that do not contain any population. This reflects the distribution of population and the fact that many areas such as airports, industrial parks and remote areas, do not have residents. For a variety of reasons, non-populated areas are sometimes designated with their own boundaries even though they have no population. "Empty" geographic units can sometimes cause surprises on computer-drawn maps when "holes" appear in the shading.

Case A: Devising National and Local Sales Territories

John Cooper has just been appointed the first national sales manager for a new line of automotive parts being introduced into Canada. These parts will be sold through an extensive network of in-house field agents calling on retail outlets. His start-up duties include creating areas of responsibility for his regional managers and helping to define districts for the local managers as well as assisting in determining the territories for the field agents.

Since John's company is very committed to market research and to evaluating sales effectiveness by closely monitoring market share changes, he must ensure that the territorial boundaries match readily available market information sources. Also, because the compensation package his company offers is quite challenging, John wants to ensure that the territorial boundaries are fair and clear. His company has a three-tier system of territories: large regions (under the direction of a regional sales manager), sales districts (managed by a sales manager) and individual sales territories (for each field agent).

Dividing Canada into Sales Regions

Based on experience in the United States, John believes that a market the size of Canada should have about six regional managers. In the U.S., his company often uses states to define the regional managers' territories. John, therefore, decides to use provinces in Canada.

John's first problem arises when he discovers that Ontario has about one third of the population of Canada and probably requires two regions, and that Quebec also might require two regions. Not being familiar with Canada, he asks his assistant to "get the right stuff to decide where to split the two larger provinces".

Provincial road maps indicate where the major cities are and John's assistant decides to split both provinces in half, from north to south. Further, being from Toronto, the assistant decides to split Ontario using Yonge Street because he knows it is Highway 11, and that it stretches all the way through northern Ontario. This split follows his boss' request for a boundary that is well understood.

John then needs to compile some basic profile data for the two Ontario regions. To get the necessary information, he visits the Statistics Canada regional reference centre in Toronto. There, he discovers that his boundaries do not coincide with the boundaries used by Statistics Canada and some other data sources such as compendia and almanacs. These documents mostly use county and municipal boundaries (CDs and CSDs). Highway 11, or Yonge Street, happens to pass through many of these CDs and CSDs. To answer his questions, John asks the data dissemination officer for assistance. The data dissemination officer explains that John can use his boundaries if he wishes to request custom tabulations for his territories. Alternatively, he could redelimit his territories. (See below for an example of his dilemma.)

John realizes that it will not be easy or practical to always request custom tabulations. Therefore, he decides to rethink his initial boundary split. Earlier, he had noticed on the road map some territorial markings and wondered what they were. The map legend says that they are county boundaries (CDs), so he asks whether Statistics Canada has a map of them as well. The data dissemination officer shows him a copy of Census Division and Census Subdivision Maps (Catalogue No. 12-572-XPB). He purchases copies of the maps for Ontario and then looks up the county populations. He writes these figures on the map as well. Finally, he picks a split that is roughly equal in total population and is easy for travel.

Dividing Regions into Sales Districts

John's next challenge is to establish sales districts within the regions. Having now completed some research on how data are published, he knows better than to arbitrarily draw lines on a map or to blindly follow major roads.

John quickly decides to use Statistics Canada boundaries to define his districts. The concept of being market-centred appealed, so he decides to start with census metropolitan areas (CMAs) and census agglomerations (CAs). These areas are Statistics Canada's delimitation of the market influence of urban areas and are formed by grouping municipalities. However, since CMAs and CAs do not cover the entire province, he needs to fill in the gaps between urban markets by assigning the remaining municipalities and census divisions. The maps of census divisions and census subdivisions mentioned earlier are very helpful in this exercise because they also show the CMA and CA boundaries as well as all remaining municipalities.

Dividing Sales Districts into Individual Sales Territories

The next step is to subdivide major urban centres into local field agent territories. John has a number of options: census tracts (CTs), his own designations, or postal code forward sortation areas (FSAs). Since he would not be working very much with individual customer list information, he decides not to follow FSAs (see Box at bottom of page). The choices remaining are to work with established CT boundaries or to create his own boundaries. If he creates his own boundaries, he would need to ask Statistics Canada to compile the census data for them (or an authorized census distributor). John decides on CTs for two very pragmatic reasons: it is cheaper and he could avoid disagreements about why a boundary went where it did, by saying he followed established boundaries.

Each CMA and tracted CA (CAs with an urban core of at least 50,000 in a previous census have census tracts) have a special set of profile publications that summarize a wide range of census data in each area (colloquially known as the "CT Profiles series"). These publications include maps of the CMA's or CA's census tracts.

Additionally, there is a compendium of census tract maps called Census Metropolitan Areas, Census Agglomerations and Census Tracts Maps. John uses the latter because he wants to analyse a number of places. He followed the same process of writing population counts on the maps and then adding census tracts together into groupings of equal populations. In practice, this can be quite a tedious process if large urban areas are being divided into a number of local sales territories. One alternative is to approach a research company with the capability to do this allocation using a computer system. This way, a number of other factors could be weighed such as income and family status. Another option is to purchase a mapping software package that would be used in-house to do the same thing (see Case E). The arithmetic of adding up the CT populations could be simplified by purchasing the data from Statistics Canada on diskette for importing into a standard spreadsheet program or by using GIS software with the census tract digital boundary file or digital cartographic file.

What if it is too late, boundaries are already drawn up, and they don't follow CTs or other census geographic units? See Case C.

Postal Codes or Census Geography

For most organizations, a key question that arises when defining territories is whether to use postal geography or census geography. If the primary source(s) of information will be address list(s) such as customer files and census data will only be used periodically with these data, then it is probably best to use postal geography to set territories; more specifically, to use FSAs. However, if postal sources are not involved, then using census geography will probably make more sense. There are ways of converting data from one basis to the other such as with the help of Statistics Canada's Postal Code Conversion File. However, such conversions are not always exact so that the base system chosen should reflect the most important source of information being analysed.

Case B: Evaluating the Local Market Around a Site for a Store

A young entrepreneur, Renée St. Jacques, has obtained the rights to open a new franchise in a chain of supply stores serving home-based businesses. She has two sites in mind in Sherbrooke and wants help in evaluating the market around each site. The head office of the franchise has given her a little information on target markets and typical customer travel patterns. With this in mind and no experience in retail site location research, she drops into the Statistics Canada regional reference centre located in Montréal, looking for the required "demographics".

Renée was keen to learn and to do the research herself. In thinking through what she needed and in reviewing some of the material given to her by the franchise head office, she decided she needed to determine the following:

(a) what the trading area of a store would be. The trading area is that geographic area around a site from which the majority of customers would come, considering transportation routes, competitor locations and the customers' willingness to travel certain distances;
(b) the demographic description of her target group who were or could become home-based businesses (provided by the franchisor);
(c) how many potential clients would be in the trading area of the sites (to be calculated), and whether this met the threshold size recommended by the franchisor.

The franchise's head office explained that the normal trading area for her store was a few kilometres, depending on transportation and other physical factors. She bought a street map of Sherbrooke and traced out a rough guess at what the trading area would be. She took into consideration the river and the bridge that squeezed and stretched her trading area. She took the map with her to the Montréal Regional Reference Centre.

In discussions with the staff at the Regional Reference Centre, she quickly discovered that the most common way to geographically define market areas within larger urban cities in Canada is to use census tracts. These are statistical units of geography intended to divide large cities into smaller neighbourhood-like areas, and for which considerable census data are produced and can be easily obtained.

She was shown a map of census tracts for Sherbrooke and compared it with her own drawings on the street map. Without too much difficulty (see mapped example below) she was able to determine which census tracts most closely resembled her drawing. In some cases, the fit was not very exact, and she had to make a decision on whether or not to include the particular census tract. She did this with both sites and ended up with a list of census tracts for each site that approximated her two likely trading areas. It was then a simple matter to pull out the demographic data for those census tracts from profile publications from both the 1996 Census and the 1991 Census. Since census tract boundaries are generally held constant from census to census, it was easy to compile comparable data from both time periods.

In further discussing her data request with Statistics Canada, she found that she could have attempted to follow more exactly her trading area boundary by retrieving data tabulated at a more detailed level of geography (see Box below). This would normally be a computerized process much along the lines described in the second part of Case C presented next.

Tips in More Complicated Situations

If many census tracts (or many sites) are involved, Statistics Canada provides a service to automatically add the data together. Or, electronic data files could be purchased and the user could do the work in-house.

Closer approximations of exact trading area boundaries can be accomplished using data tabulated at the enumeration area, or even block-face level of detail in the larger urban centres. This is warranted if trading areas are small – less than a kilometre or two in radius.

Case C: Custom Designing Your Own Boundaries

Jill Stasiuk is a planner in a provincial department of municipal affairs. She also sits on an interministerial task force examining a number of regional planning issues. As an accomplished regional analyst, she is very familiar with a wide variety of analytical tools and regional data sources, and often gets asked for help on complex research questions. This week she has two: what is the population of a specific segment of a township that is probably going to be annexed to a neighbouring town; and, what are the population trends in a series of health districts in the southern half of her province.

A Special Population Compilation Request

Jill dealt with the annexation question first. A map supplied by a colleague outlined the area under consideration for annexation. She had a number of choices for calculating the population of the area. She could ask the group responsible for maintaining assessment roles to tap into their database for the properties included in the study area. In her province, this database also includes a list of inhabitants and their ages. Also, she could commission a survey of her own in which someone would canvass the area and get the information directly. The last option was

that she could look up the latest census data. All three had their limitations in terms of accuracy (both assessment and census being a couple of years old) and cost (a field survey being expensive). She decided to get assessment data and census data first before recommending a survey.

The census data approach became complicated. A comparison of the map provided by the colleague and census enumeration area maps for the area revealed that the annexation actually split a number of EAs. In a couple of instances, the boundary followed streets in a built-up area, and although they split EAs, Jill knew that Statistics Canada would also have coded the data to the block-face (one side of a street between two intersections) and could easily retrieve the data. In fact, her ministry had the Block-face Data File that was released for the first time by Statistics Canada following the 1991 Census. Therefore, she could do the calculations herself with the appropriate software and geographic reference files for the urban portion of the annexation. However, the annexation also included some rural territory that was not included in the block-face program.

At this point she contacted Statistics Canada's Regional Reference Centre to request the custom population compilation service. For this service, Statistics Canada staff go through the original 1996 Census visitation records of the census representatives and identify which households are where. This is a manual process and must be done by Statistics Canada staff to protect the confidentiality of respondents.

Drawing Your Own District Boundaries

The second request about the health districts was somewhat similar because it turned out that the boundaries had been drawn without regard to municipal or census geography, but had been done to minimize distance from hospitals and clinics. Thus, standard published census tabulations were not going to work. Also, because there were quite a number of large districts, it was impractical to manually look up on many maps which EAs were where.

In discussions with Statistics Canada, she found out that she could request custom census data retrievals for geographic areas that did not conform to census geography. The first step was to mark **clearly** the health district boundaries and their names on maps supplied by Statistics Canada. At first, she had supplied a regular provincial road map on which she had drawn the health districts using a magic marker. However, this was too rough since the line was wider than some towns it passed over on the map, leaving it unclear whether or not to include them. To avoid this type of problem, Statistics Canada supplies base maps to its clients.

The health district boundaries are then "digitized" (see Box) from the map to create a "digital" boundary file. Statistics Canada then produces plots of the boundaries and population counts for each health district for Jill to review before the census data are retrieved. Pending her approval of the digitized boundaries, the required data are then extracted.

"Geocoding and Digitizing"

"Geocoding" is the technique used to geographically code and link households to small geographic units in support of data retrieval. For instance, customer addresses can be geocoded by matching their postal code to Statistics Canada geography using the Postal Code Conversion File. Thus, customer address files can be tabulated according to the census geography, and the corresponding census data examined for the demographics of those areas in which the customers live.

Statistics Canada geocodes households to a block-face representative point in large urban areas (where Statistics Canada maintains computer files of the streets and other network features – see the ***Street Network Files*** concept), or an enumeration area (EA) representative point in small urban and rural areas. This links all the census data for the households to a particular EA or block-face representative point.

"Digitizing" is the process of converting map data from their original paper form to a digital format. This is the first step in a retrieval of statistical data for non-standard areas. Then computer processing is used to retrieve data for that area. This is done by calculating whether the representative point is inside or outside the digitized boundary (see illustration below). If the representative point(s) is inside, then all data for the EA or block-face are included. **Note**: To each block-face (within Street Network File coverage) and enumeration area, the census data for that area are linked. It is this complex linking process that permits the extraction of data for non-standard as well as standard areas.

Case D: Using Direct Marketing to Promote a Retail Operation

Kim and Carol Lee own an upscale children's toy store in a Vancouver area shopping mall. Over the years, they have built up a loyal customer base from the neighbourhood surrounding the mall. They would like to expand their operations by adding a mail-order service. They are strong believers in promotion and would like to use direct mail to promote their store in the areas nearby. Carol decides to take charge of the planning.

After a strategic review with an advertising agency, Carol opts for a flyer drop in the surrounding area to promote her store.

Planning a Flyer Drop Around a Store

Carol decides to do a flyer drop around her store to take advantage of the low cost per advertising piece of unaddressed mail, even though such mail is less likely to capture the attention of the resident than is addressed mail. Carol's first call is to the post office to find out what their guidelines are for flyer distribution. Her second call is to the mall to find out if other stores at the mall have done similar flyer campaigns. She finds that the gift store has done a Christmas campaign but that they used a flyer distribution company and not the post office. After comparing prices and other considerations, Carol elects to use the post office because there are a lot of apartment buildings in her area and she feels the post office will do a better job of delivering the flyers to each apartment mailbox.

In her discussions with Canada Post's Admail representative, Carol is asked to decide what neighbourhoods she wants to cover. She is shown a map of FSAs (forward sortation areas are mail districts identified by the first three digits of the postal code). FSAs are postal delivery territories.

Carol was then asked to specify which FSAs she would like to include. Distance is one criteria she uses that she combines with the idea of demographically targeting certain areas. She knows through experience that her store appeals to grandparents as much as parents and children. Given the upscale nature of her products, she decides that some local neighbourhoods will be better than others based on age, income and family status. In order to pick the best FSAs, she consults a map to pick the closest geographic alternatives. Since she has some choices to make as to whether to include some FSAs or not, she then decides to take a look at their demographic profiles using the FSA Profile Series from the 1996 Census. The choice became a tradeoff between costs per piece, likely success rates and her budget. Over time, Carol was able to refine the list of FSAs she included based on experience.

One basis for the refinement to the planning, was to research the true extent of the store's trading area. This is the area that a majority of their customers come from. While Carol and Kim had a rough idea, they decided to test their belief with a simple in-store research test. To do this, they put a map of their part of town up on a wall and asked each customer that came in over the next month to put in a pin indicating where he/she lived. At the end of the process, it was easy to then draw a boundary around the areas of highest customer concentration and refine the choice of FSAs accordingly.

What's the Alternative to "Pin Maps"?

While a "pin map" for one neighbourhood store is easy to organize, many situations arise in which it becomes impractical. More computer-literate marketers can take the "pin map" approach using a computer mapping program and Statistics Canada's Postal Code Conversion File (PCCF). This file contains a set of geographic coordinates for six-digit postal codes in Canada. Therefore, after matching to the PCCF, a list of customers by postal code can be fed into a computer mapping program, which will plot customer locations based on postal codes. Alternatively, Statistics Canada or a number of secondary distributors will provide this custom service upon request.

Case E: Setting Up Your Own Computer Mapping System

Jim Thompson has just started a new job with a large retail chain in their real estate department. The company recently purchased some mapping software to help plan delivery routes and he wondered whether he could use it to plan store location and store marketing campaigns. He had previously worked for a market research company that had used a variety of systems, and thus knew a fair deal about the applications, but never had "built" a system from the ground up.

In reading the documentation for the software purchased by his new employer, Jim found that it had many mapping and analysis capabilities. From the literature and his previous experience, he knew that mapping software was great for displaying information and for analysing many different locational questions. His new package seemed to have all the features he wanted and he was keen to get started with it.

Because the original purpose of the purchase was to determine the best routes for delivery, the only geographic information stored in the system was a set of street files that had been obtained from Statistics Canada: the Street Network File (SNF). This is a very comprehensive computer file covering the streets and key physical features of most larger Canadian cities, including the urban portions of CMAs and larger CAs.

However, Jim needed to be able to portray areas for which census data are published, such as municipalities and census tracts. He found that while he could draw his own boundaries, he could not produce a map of Canada showing counties (census divisions) since he did not already have the boundary files and he did not wish to draw all 288 boundaries himself.

In order to plot anything geographically, digital boundary files are needed. These are computer encoded coordinates that allow dots and lines to be traced out and for information relating to them to be also displayed. Each software package has its own computer format for these instructions. Thus, Jim knew that he would have to make certain any boundary files he purchased could be transformed by his system if they did not already automatically conform to his software.

Jim's first decision is which boundary files to purchase. Since he knew he would be doing extensive analyses with census data, he realized he would need to display much of the census geography.

What Computer Mapping and Geographic Analysis Systems Can Do for Demographic Research

A wide number of capabilities are now available including:

- distribution maps of census data and customers using dot maps, shaded maps (choropleth maps), pillar maps, 3-D maps, etc.;
- determining the extent of trading areas covering x% of customers;
- calculating and drawing the most efficient territory alignments taking into consideration the locations of outlets, competitors and customers;
- overlaying different distributions and geographic features;
- combining demographic models, such as population projection equations, with visual presentations of output;
- combining statistical analyses with graphical output such as pie charts and scatterplots;
- retrieving and computing demographic data for custom-drawn areas (either based on digitized boundaries or drawn right on the screen); and
- computing distances, drawing routes and calculating densities.

You should research the capabilities of your own mapping or Geographic Information System (GIS) software package since they vary greatly.

A quick call to his original software vendor (or he could have called Statistics Canada) told him that Statistics Canada produces digital cartographic files (DCFs) for almost all geographic units for which it releases data. The boundaries include provinces, census divisions, federal electoral districts, census subdivisions, census consolidated subdivisions, urban areas, census tracts and enumeration areas. After reviewing what historically had been analysed, Jim opted to begin with the census tract, census subdivision and census division boundaries.

Jim also realized that he needed other types of geographic files. For instance, he wanted to be able to plot customer locations. For this he needed to be able to plot the location of a street address. One tool is Statistics Canada's Postal Code Conversion File (PCCF) which includes an "x-y" co-ordinate for six-character postal codes that mapping software can use to plot location. In urban areas, this usually represents one side of a street between

two consecutive intersections. In effect, with the PCCF, a list of postal codes can be plotted on a map automatically as a series of points showing approximate location.

Jim already had the capacity to work at a very detailed level of geography due to the use of the Street Network File. He therefore decided to work with the block-face capabilities of the PCCF in urban areas. In other words he wanted to be able to plot customer locations to the block where they were located.

If Jim had started from scratch in his foray into computer mapping, he would have had to make a number of key decisions. Some mapping software comes only as a stand-alone software system and the user must acquire boundary files and data files as well. Some companies who primarily market the mapping software also assist clients with this data and boundary file process. Finally, there are companies that package the software, data and boundary files together. Other mapping and analysis systems are offered for sale as fully integrated packages of analysis, data, mapping and graphing capabilities. Only organizations that have made licensing arrangements with Statistics Canada can sell or otherwise provide census data or boundary files developed by Statistics Canada.

This page was intentionally left blank.