



ENVIRONICS

A N A L Y T I C S

**Technical Documentation
Demographic Estimates and Projections 2012**

Demographic Estimates & Projections 2012

February 2012

Source

Environics Analytics, Statistics Canada, The Centre for Spatial Economics (C4SE), Canada Mortgage and Housing Corporation, Citizenship and Immigration Canada.

Purpose

Environics Analytics has created Demographic Estimates and Projections (DEP) for an extensive set of variables at the small area level. For 2012, we release estimated numbers for the current year and projections for 2015, 2017 and 2022. In addition, the latest DEP includes historical estimates from five years before the current year (2007), which use the same advanced modelling techniques as the other estimates and projections.

Created by an innovative methodology and supplemented by government estimates and economic data like building permits and immigration statistics, DEP features authoritative estimates for a multitude of variables. We rely on the best practices of econometric forecasting based on the work of our partner, The Centre for Spatial Economics, along with demographic forecasting and geospatial estimates and projections overseen by our Chief Demographer, Dr. Doug Norris, and our Chief Methodologist, Dr. Tony Lea. Most of the calculations were made by Senior Demographic Methodologist Robert Dominico, along with Research Analysts Peter Miron and Sean Howard.

Comparability Over Time

It is common practice for data suppliers to produce new estimated and projected data annually. In this process, there is very often new reliable data from the past (e.g. from 2008 or 2010) that are used to estimate the numbers. In addition, new or improved estimation methods typically are used that differ from the past year or years. Because of these changes, it is **not recommended** that data users compare numbers (relating to any given year) from our 2012 released data with numbers that were produced in 2011 or earlier years. The comparison may show numbers going up or down for any geographic area "unexpectedly" when, in fact, the differences may be due only to different ingredients or algorithms used in the two years.

However, in each release of EA's DEP data, we provide numbers that are produced using the same ingredients and methods and these numbers can be legitimately compared. We create "trend reports" that show these results. In almost all cases, EA provides annual estimates for these data for the period 2006 to 2022 rather than just for the subset of years that are publicly released.

If you wish to compare 2011 estimates with 2012 estimates (and we acknowledge that this may be helpful in some situations), these comparisons should be done using estimates produced in the 2012 vintage release of DEP. EA does have the 2011 series numbers in the 2012 vintage DEP for all variables in DEP (except possibly for income distributions). If you need to examine these numbers for comparison, please do not use the old 2011 vintage numbers from DEP 2011. Rather, please call us at EA and we will help you with data that exist but were not formally released.

Changes From Our Last DEP

For DEP 2012, we had Statistics Canada's estimates for Census Divisions (CDs) for the years 2006 through 2010 on 2006 census geographic areas. This was very helpful data in establishing higher-level allocations and for starting trends regarding future growth.

In past years, we have used Census Subdivision (CSD) level estimates of populations and households produced by Statistics Canada along with the CD level estimates of these variables (that we have continued to use). For the preparation of DEP 2012, we undertook the estimation of the CSD level of geography ourselves. The CSD level estimates produced by Statistics Canada for historical years make the simple assumption that a given CSDs' share of its parent CD remains the same over time. While in the past we have been satisfied with this assumption, as we have moved farther from the 2006 census and alternative sources of information have become available EA has found a few important CDs where this assumption no longer holds. A prime example of this is in Peel where Brampton has grown more rapidly than Mississauga or Caledon since 2006. Thus, Brampton's share of the Peel CD has increased over time, while Mississauga and Caledon's have decreased. Our in-house produced CSD estimates for the historical years (and consequently current and projection years) better reflect the internal dynamics of many CDs across the country because we are using additional data to account for this internal dynamism. It is important to note that this methodology change does not affect the CD or higher levels of geography, but it does impact the CSD and lower level of geographies.

The primary input to this update was municipal (CSD) level housing start data and the 2006 census CSD shares of their parent CD. Additionally, the 2001 to 2006 rate of change in share by CSD and taxfiler data (produced by Statistics Canada Income division, formerly known as SAAD) were also used as secondary inputs. We computed the change in share of housing starts for each CSD in a CD from 2006 to 2011. These annual rates of change were applied to the 2006 census shares for 2007 to 2011. Beyond 2011, the 2006 to 2011 trend was projected forward using proprietary techniques. We were able to obtain housing start data for all CSDs in CMAs and CAs. Where housing starts data did not exist (rural Canada), we relied on the 2001 to 2006 change and taxfiler data and where that data was deemed unreliable we reverted back to the simple assumption of constant shares among the constituent CSDs for a CD.

Levels of Geography and Years for Data Release

This document relates to Environics Analytics' 2012 vintage DEP for census and census-type demographic and socioeconomic variables. These are created for every level of 2006 "census geography" down to the DA. They are also available for all but the smallest Forward Sortation Area Local Delivery Unit (FSALDU) levels of postal areas nationally. They can be linked to FSALDUs expressed as counts, percentages, averages, etc. via EA's Enhanced Postal Code Conversion File (PCCF). "Estimates" are for the periods between the Canadian Census 2006 and the present. "Projections" are for beyond the present. EA projects its variables every year from 2012 to 10 years into the future. However, EA releases only the following years as part its off-the-shelf offering:

- 1) current year estimates (2012)
- 2) 5 years before the current year (2007)
- 3) 3 years beyond the current year (2015)
- 4) 5 years beyond the current year (2017)
- 5) 10 years beyond the current year (2022)

This year is the fifth year that we have also undertaken estimates for 5 years back from the current year, now to 2007. This is a convention that we will continue in future releases that will create an historical estimate that uses the same general methods of the other years, thereby creating a comparable historical value.

In DEP, we actually project each year from the last census to 10 years beyond the year of DEP—even though we have a subset of years that are part of our standard release package. If clients require years that are not part of the package, these data can be assembled; clients should contact their sales representative.

Centre for Spatial Economics

As part of our DEP process, EA retained the services of the Centre for Spatial Economics (C₄SE) and Tom McCormack, its President and Principal of Strategic Projections Inc., to undertake estimates and projections of all key base variables at the level of the CD and above for all of Canada, and from the current year (2012) to 2022. These variables included population and household counts, labour force variables, immigrants by country of origin, average household incomes. In the past, we also received CD-level household income distributions, but we now make these estimates on our own. With few exceptions, EA used the C₄SE projections as the target in undertaking small area estimates and projections; except where specifically noted, EA's DEP variable values are consistent with those of C₄SE.

Census Coverage Error Adjustments

These estimates and projections adjust the census base year numbers to reflect the census undercount, and sometimes overcount, factoring in the net coverage error that is known to exist. Thus, all of the EA estimates and projections factor in this adjustment. This means that **these numbers are not comparable with census data** which has not been similarly "corrected". The EA correction was first estimated and projected based on information released by Statistics Canada and C₄SE at the CSD level using a process that "controlled them up to" more reliable statistical estimates at the provincial level. The coverage error factors were then estimated for the DA level by EA using proprietary methods, and then controlled up to the CSD level. This means that the aggregation of the DAs in any CSD has been designed and computed so that they match the CSD-level estimate calculated in advance.

Raking

The expression of raking, or raked, is a technical term in statistical analysis. As used below, it refers to a process of systematically adjusting numbers for the value of a variable defined for constituent areas, or other "subsets", so that these numbers add up to a prior estimate or projection considered to be more reliable for a larger area or set containing all of those constituent areas/subsets. The expression "controlled up to" is a synonym for "raked to" as in "... EA's DA level estimates for these variables were then controlled up to the CSD-level numbers." Raking is necessary to ensure that all estimates and projections are consistent and can be added in a logical manner.

Sprinkling

Sprinkling is the name given to a process of creating integer (count) data when it was derived first in a non-integer format. In general, most of DEP is first done in non-integer format as this is much more straightforward. Then the counts are sprinkled to create the final deliverable. Sprinkling is more than rounding because sprinkling requires that the final data set being sprinkled sums to exactly the right control total (for a level of geography). It also must often add up to a known figure from a more general data set (persons as opposed to females). We have developed a very good sprinkling algorithm that is used in all cases. The process rounds up (down) those records that have the highest (lowest) decimals and are also the most likely to have a higher (lower) next integer value.

Overview of Variables in the 2012 DEP Release

The following is an overview of the variables that have been estimated and projected in this 2012 release of Environics Analytics demographic data, and some high level notes on methodology.

1) HOUSEHOLDS AND POPULATION

a) Definition

- i) A "household" is a private dwelling occupied by one or more persons.
- ii) The "population" universe includes all persons whose usual place of residence is in Canada. This includes citizens, persons with landed immigrant status and non-permanent residents such as students, persons with work permits and refugee claimants. This group is referred to as "total population." In addition to total population, we also worked with a subset base population called "population in private households" (see b)iii) below). We also refer to this

- group as the "household population." It is important to pay attention to which of these bases is used; the failure to do so may result in errors of analysis.
- iii) Estimates and projections of population are derived from a wide range of data including household estimates discussed above, the average household size from the 2006 Census and temporal trends in these rates, population estimates and projections for Census Divisions from C₄SE and Statistics Canada Annual Demographic Statistics for provinces and Census Divisions
- b) Methodological Issues
- i) The reference counts for estimates and projections at the DA level come from the households' counts from the 2006 Census adjusted for the census "undercount" or what is more appropriately called "the net coverage error." Net coverage error is an estimate done by Statistics Canada of the number of persons not enumerated or enumerated more than once. Statistics Canada released estimates of the net coverage errors at the national, provincial, CD and Census Metropolitan Area (CMA) levels, and these served as the basis for our coverage error estimates for 2006. The EA household estimates were controlled to the EA proprietary C₄SE 2006 estimates by age and sex at the CSD level. Henceforth, the coverage error correction was simply continued.
 - ii) Estimates and projections of households are derived from the base year counts using a proprietary methodology with inputs from Census data, patterns of household and population change at the DA, CT and CSD levels, and population and household estimates and projections for CSDs from C₄SE. Proprietary data and methods were used to identify small geographical areas undergoing growth or decline and how much change was occurring.
 - iii) The total population has two high-level subgroups: the population in private households ("household population" henceforth) and the population in collective dwellings ("collective dwellings population" or simply "collective population"). Collective dwellings include work camps, residential religious institutions, hospitals, some seniors' homes with more collective services, rooming houses, prisons, and university residences. Note that the Census collects very little information on the collective population.
- c) Benchmarks
- i) "Total population" and "household population" are two common bases, benchmarks, or universes for other population-based variable groups. Household population as indicated above refers to the population in private households. Topics relating to population in EA's DEP are based on either total population or household population. These two bases have both been used by EA because some data were produced by Statistics Canada on one base and not the other. We will state which base was used in each case.
 - ii) The total population and household population bases or universes have a number of sub-universes.
 - iii) "Population 15 or over" refers to all persons in the total population universe who are in this age group derived from the distribution of five-year age groups.
 - iv) "Household population 15 or over" was derived from "total population" 15 or over less a specified percentage of the collective dwellings' population designed to estimate the collective population 15 or over. The percentage was derived from census data and trends over time in the age and sex of persons in collective dwellings.
- d) 2001-2006 Walkover
- i) The 2009 and later-vintage DEP is released for 2006 census geography. Changes in population and households counts for 2001 to 2006 were used as part of our whole estimates and projections process and therefore all of the 2001 census inputs were rebased or walked over to 2006 census geography at the DA level. This was done using carefully constructed DA2001 to DA2006 proprietary correspondence files created by EA.
 - ii) For more information on the walkover process, see the supplementary document called "**Walkover Methodology for Census 2001 to/from Census 2006**" which is available from Tom Montpool.

1) POPULATION BY SEX, AGE AND AGE AND SEX

- a) The universe for these variables is "total population." The variables are estimated counts by sex, by five-year age groups up to 85 or over, and by both sex and age. Total population was obtained by adding the counts of both sexes over all age groups. The 2006 census age and sex distributions were used as the initial distributions. Census subdivision age and sex estimates and projections from C₄SE were used as the main controls for DA level distributions. C₄SE took into account both macroeconomic growth factors and demographic trends in undertaking its population estimates and projections. The base economic model was a full econometric model of the Canadian economy specified by province. The base demographic model was a cohort survival model with birth rates, age-specific death rates, survival probabilities and in- and out-migration – with all of the data disaggregated spatially. Households were subsequently derived using household headship rates relating to prior population estimates or projections.
- b) Median ages for total population, females and males represent the middle value of the distribution and are calculated from five-year age cohorts. They are expressed in a commonly used format of age in years to one decimal place.

2) MOTHER TONGUE

- a) Mother tongue is the first language learned at home in childhood and still understood by the individual. The universe for these variables is household population.
- b) Variables include 44 single response languages or multiple responses for English and French. Multiple responses occur when persons have indicated that they speak two or more languages, for example English and French or Spanish and Italian. All responses other than those listed are included in the "other languages" category.
- c) The 2006 census mother tongue distribution was used to start the series and then the annualized rate of change between the DA mother tongue language numbers for the 2001 census and the 2006 census were used to generate the subsequent estimates and projections.

3) IMMIGRANT STATUS AND PLACE OF BIRTH

- a) The universe for these variables is total household population.
- b) This variable is first sub-divided into immigrants and non-immigrants; each of these is then divided into specified sub-groups. Within non-immigrants are those born in their current province of residence and those born outside their current province of residence. For immigrants, 47 countries of place a birth are given. All other countries are grouped under the "other countries" classification. The 2006 census immigrant status by place of birth was used for the initial distribution. Subsequent years were projected using trends' input from C₄SE for immigrants at the CSD level and changes in immigrants at the DA and CSD levels between the 2001 and 2006 censuses.

4) EDUCATIONAL ATTAINMENT

- a) The variables here relate to the highest certificate diploma or degree completed by individuals. There are 13 classifications that we collapsed into six categories of educational attainment variables, and they include: none, high school, college certificate or diploma and university certificate or degree. The universe for these variables is "household population 15 years or older." (Note that in the 2001 census the base was "household population age 20 plus" and not "household population age 15 plus" as in previous the census.)
- b) In the 2006 Census, the highest level of schooling distributions at the DA level were used as the initial distribution. A series of proprietary models based on multinomial logit models was used to project these distributions forward in time.

5) LEGAL MARITAL STATUS

- a) Legal marital status refers to a person's conjugal status under the law. This includes married and not married – with the latter classified in four groups. Persons in common-law relationships are

not dealt with as "common-law" in this theme because there is no such class. Instead, people who are not currently married fall into one of these four classes: never married, separated, widowed or divorced). The common-law status of people is dealt with under the common-law theme below.

- b) The universe for (legal) marital status is "total population 15 years or over."
- c) The variables here are legally married, single (never married), separated, widowed or divorced.

6) COMMON-LAW STATUS

- a) Living in Common-law is defined as two people of the opposite sex, or of the same sex, who live together as a couple, but who are not legally married to each other.
- b) The universe for common-law status is total population 15 years or over.
- c) The 2006 census percentage distribution over the classes was used for 2006. Other years beyond 2006 were projected based on analyses of trends estimated by EA.
- d) The variables here are living "in common-law", and "not in common-law", relationships.

7) CENSUS FAMILIES BY STRUCTURE AND NUMBER OF CHILDREN

Note: Also see the next header 8) for these family variables transformed onto the household universe.

- a) The universe for "census families by structure" and "number of children" is "census families in private households." A "census family" is defined as a married couple (with or without children of either or both spouses), a couple living in a common-law relationship (with or without children of either or both partners) or a lone parent of any marital status with at least one child living in the same dwelling. A couple living in common-law may be of the opposite or same sex.
- b) Married couples and common-law couples are added together to create "couple families."
- c) Children in a census "family" are never married sons or daughters living at home, children previously married and now living at home with one or more parents but with no spouse or common-law partner, and grandchildren living with their grandparent(s) but with no parents present.
- d) "Family population" is the total number of persons in "census families."
- e) "Average children per census family" is the total number of children living at home divided by the total number of "census families."
- f) The universe for "families by structure" and "number of children" is "census families in private households."
- g) The variables in this set include those relating to "census families": "without children at home," "with children at home" and this "children at home" group broken into "1 child," "2 Children" and "3 or More Children." This set of five variables exists for each of these groups:
 - couple families
 - married couple
 - common-law couples
 - lone parent families
 - female headed
 - male headed

The set ends with the variable number of family persons.

8) CENSUS FAMILIES BY STRUCTURE AND NUMBER OF CHILDREN TRANSFORMED ONTO THE HOUSEHOLD UNIVERSE

This set of variables is a special recalibration of the set of census family-based variables described in the last heading. For details on the variables involved in this set, please see heading number 8. Since the variables here were derived first from the family universe, the methods of section 8 had to be applied first.

- a) Explanation of the Need for a Household Universe Version
 - i) For most general marketing and targeting applications, this household version of the data will be the appropriate one to use. The reason for the adjustment is that, for many purposes, the

- reference denominator for target marketing is households. Families' counts differ from households and are frequently very different. There are two reasons for the difference:
- (1) single persons living alone are households but not families, and do not qualify as "family households"
 - (2) some households have more than one family living in them – multiple family households.
- ii) The universe for this variable is "census family households" (see also topic 21).
 - iii) If a marketer wished to target "couple families with children" (or female headed lone parent families with two children) and wanted to compute key statistics, it is important to be aware that there are two possible approaches, though one is normally inappropriate. If the marketer had a list of names and addresses or all points on a map representing census **families**, then "families" is the best benchmark or denominator. One should compute the target group divided by total census families or a subset of families. If this were a DA with 100 census families and 20 in the target, then the "penetration rate" would be 20%. (This relates to all census families and it does not deal with the issue that this DA may have only 80 households.) But it is uncommon for marketers to have a reference list or map of families. Normally, marketers are looking for target groups of families among all households. Indeed, much software in marketing makes things seemingly "simple" for marketers by assuming that the reference denominator is households and not families.
 - iv) In the situation in which marketers are going to be using family-based variables as if the counts related to "households" as opposed to "families," then it is helpful to have the **families data** adjusted or rebased to a household universe. This means that the adjustment process has to deal with multiple family households.
 - v) Let us make an example database:
 - 100 families
 - 80 households
 - 60 households of single family households
 - 10 households of multi-family households
 - 10 households of non-family households (e.g. singles)
 - 70 family households (60 with single families and 10 with multiple families)
 - The analysis has to take 100 **families** and transform them into 70 **family households**. Sixty of these have been transformed already because they are already single-family households. Some judgment is required here about how we should transform the aggregated families in the multifamily households. The example data above will be used in further discussion below.
 - The variables in this section that we have created are simply the adjusted versions of the counts under the family variable heading 7) above. In a situation in which there are no multifamily households, the counts will be the same as in the previous unadjusted family data. Unlike all other DEP and census variables, we have left the estimates and projections data here in non-integer form. This is because these family households are a re-conceptualization of non-existent households and there is no reason that they should be integers logically. One reason for leaving them in this decimalized form is so that users realize they are using an artificial construction.
 - vi) For more on the topic of transforming the family universe data to a household universe, please see the supplementary document entitled "**Transforming Family Universe Data to the Household Universe**" which is available from Tom Montpool. It has a more detailed example to help explain the methodology.

9) LABOUR FORCE ACTIVITY

- a) "Labour Force" refers to persons aged 15 years or over who were either employed or unemployed. Unemployed persons are defined as those who have no paid work at present but

are actively seeking employment. Persons with no paid work who are not actively looking are considered to be "not in the labour force" (as opposed to unemployed).

- b) For these estimates and projections, the participation rate is defined as "persons in labour force" divided by "household population 15 or over."
- c) The universe that EA used for labour force is "household population 15 years or over." Census variables for labour force data are published for a slightly larger universe of persons 15 years or over "who are not in institutions". For methodological reasons, the EA universe for estimates and projections additionally excludes persons who reside in non-institutional collective dwellings.
- d) Variables include: "persons in the labour force," "labour force participation rate" and "unemployment rate."

10)HOUSEHOLD MAINTAINERS - BY AGE

- a) The universe for "household maintainers by age" is private households. The maintainer is a person or persons in the household who pay the rent or the mortgage, taxes or electricity, etc., for the dwelling. There can be more than one maintainer. For "maintainer by age," the age of the "primary maintainer" is used.
- b) "Median age for household maintainer" represents the middle value of the distribution and is calculated from 10-year age cohort distributions. It is expressed in a commonly used format of age in years to one decimal point.
- c) The 2006 census distribution for ages of household maintainers was used as the base and C₄SE estimates and projections were used for post-2006 years in a relative manner.

11)HOUSEHOLDS - BY SIZE

- a) The universes for "households by size" are "private households" and "persons in private households." Household size refers to the number of persons occupying a private household.
- b) Variables for "households by size" include: "households with one person," "households with two persons," "households with 3 persons," "households with four or five persons" and "households with 6 or more persons."
- c) The 2006 census percentage distribution was used for all years' estimates and projections as there are no more recent data trends that can be used here.
- d) The final step in the creation of "households by size" at the DA level was to make minor adjustments to enforce a control that the distribution of households over the size classes was consistent with "average persons per household" estimates undertaken in an earlier step.

12)STRUCTURAL TYPE OF DWELLING

- a) The universe for "structural type of dwelling" is "total occupied private dwellings" (also referred to as "private households").
- b) The variables included here are: "occupied private dwellings" (usually referred to as dwellings) by classes of structural characteristics and/or dwelling configuration, that is, whether the dwelling is a single-detached house, a single attached house, a row house, a duplex, an "other single attached house," an apartment in a building with 5 or more storeys, an apartment in a building with less than 5 storeys, or a mobile home.
- c) The 2006 census percentage distribution was used for all years' estimates and projections as there are no more recent data trends that can be used here.

13)HOUSEHOLD INCOME

- a) The universe for income is "private households."
- b) The variables included here are households by income classes of \$10,000 ranges, up to the class "\$100,000 plus," "average household income," "median household income" and "aggregate household income." A very significant effort was made to ensure that the very best assumptions and rigour were used in this estimation and projection exercise because income estimates are so important to most users. We release income projections in two units of measurement: "nominal or current dollars" and "constant or real dollars." "Nominal or current dollars" have not been

adjusted for inflation, while those that have been adjusted for inflation are called "constant or real dollars." While current dollars reflect the face value of a dollar for the year in question, constant dollars reflect the purchasing power of that dollar according to a base year's value. The base year for constant dollars is 2005. A separate document is available for those wishing a more detailed description of the differences between these concepts and an indication of when to use each. The "**Constant Dollars Versus Current Dollars**" document can be obtained through your EA account representative or from Tom Montpool.

- c) Average household incomes were estimated first. A set of control total estimates of average household incomes were undertaken at the CD level by C₄SE using an econometric/demographic model. These CD level estimates were used as controls for a CSD level average household income econometric model. First, the 2005 incomes (the incomes released in the 2006 census) of those CSDs that had been suppressed by Statistics Canada were estimated based on using other unsuppressed income-related variables and taking into account incomes of unsuppressed similar neighbours. The 2005 incomes of many suppressed DAs were estimated using similar methods. Special econometric algorithms were then used to project the average household incomes of DAs and the income distributions forward one year at a time so that areas with the types of households most likely to enjoy faster and more consistent increases in income had their income grow faster. This dynamism was tempered by exploiting relationships found in past trends and raking the DA-level averages in a non-linear manner to the CSD average household incomes. Where possible CD and CSD incomes were controlled to published historical trends from the Canada Revenue Agency.
- d) Aggregate household incomes were computed after average household incomes by simply multiplying the projected number of households by the projected average household incomes at the DA level. For other levels of geography, the DA aggregate income value was summed up, divided by the appropriate household count and rounded to create the higher level averages.
 - i) The initial DA-level averages are unrounded (e.g. they have decimal places as cents). It is unreasonable to release such seemingly precise data so the release database has values as integers. The DA aggregates could also have been calculated using the unrounded averages but this would result in numbers that may not make sense to clients who would naturally take the rounded household counts and the rounded average and get a result that didn't match the released aggregate.
- e) Income distributions are done for the following 11 classes: under \$10,000, \$10,000 to \$19,999, \$20,000 to \$29,999, \$30,000 to \$39,999, \$40,000 to \$49,999, \$50,000 to \$59,999, \$60,000 to \$69,999, \$70,000 to \$79,999, \$80,000 to \$89,999, \$90,000 to \$99,999 and \$100,000 and over. Adjusting income distributions is a very complex process. Income distributions were projected after the average household incomes were projected. The projection series went from 2005 through 2022. If the base year (2005), household income distributions were suppressed (missing), we first estimated the missing average household incomes and then created the household income distributions to correspond with the new averages. Different multivariate statistical methods were used for these tasks. After the missing data was estimated for both the average household incomes and household income distributions, a time series approach was implemented to project change in both the average and distributional characteristics of the household incomes from higher levels of geography to lower levels. As a key ingredient, we used C₄SE's average household income projections for all CDs to 2022. Then, a time series projection was applied to the average household income of the DAs to ensure that their means were consistent with the CSD means. Then we did projections of income distributions for CSDs using non-linear categorical predictive models so that the implied means of the distributions were consistent with the previously projected CSD means. Finally, we projected the DA income distributions so that the distributions were consistent with previously projected means and the CSD distributions so that they added up to the CSD projections in each year. The DA income distributions were projected within a large non-linear mathematical programming system which ensured that all constraints were satisfied and maintained the shape of the past income

distribution while nudging it towards a new appropriate equilibrium and softened temporal distortions.

- f) In early releases, the difficulty of the optimization-based projection limited EA estimates of income distributions to the current year and the five-year projection. However, this limitation was overcome in the 2009 and subsequent releases which include incomes distributions for all release years including 2006. Income distributions for other years are available by special request only.

14) HOURS SPENT LOOKING AFTER CHILDREN – BY SEX

- a) The universe for "hours spent looking after children" is "household population."
- b) The variables include "persons by number of hours spent looking after children, without pay."
- c) The 2006 census percentage distribution was used for all years.

15) HOURS SPENT CARING FOR SENIORS – BY SEX

- a) The universe for "hours spent caring for seniors" is "household population."
- b) The variables include "persons by number of hours spent caring for seniors, without pay."
- c) The 2006 census percentage distribution was used for all years.

16) VISIBLE MINORITY

- a) The universe for "visible minority" is "household population."
- b) The variables include the thirteen major visible minority groups defined by Statistics Canada. One of the classes is "having no visible minority."
- c) The 2006 census percentage distribution was used to start the series. Then the annualized rate of change between the DA's visible minority numbers for the 2001 census and the 2006 census was used to generate the subsequent estimates and projections.

17) OCCUPATION

- a) The universe for "occupation" is "total labour force."
- b) The variables include the major occupational groups from "national occupation classification" (NOC) and a "not applicable" classification for "persons in the labour force with no occupation."
- c) The 2006 census percentage distribution was used for all years.

18) PERIOD OF CONSTRUCTION OF DWELLING

- a) The universe for "period of construction of dwelling" is "occupied private dwellings."
- b) The variables include construction time periods from before 1946 to after 2006.
- c) The 2006 census distribution was used as the base for DEP 2012 and for 2006. Subsequent years include special estimates for the group "post 2006" based on additional data.

19) TRAVEL TO WORK

- a) The universe for "travel to work" in DEP 2012 is the special "travel to work labour force." This is the reference to people who have a regular place of work that is not at home and includes those like salespersons who work out of their cars. (In the past, up to DEP 2010, the universe here was "total labour force.")
- b) Variables include the different modes of transportation to work such as car, public transit, etc.
- c) The 2006 census percentage distribution for travel to work was used for all years.

20) NUMBER OF CHILDREN AT HOME

- a) The universe for "number of children at home" is "total children."
- b) Children in a census family are never married sons or daughters living at home, children previously married and now living at home with one or more parents but with no spouse or common-law partner, and grandchildren living with their grandparent(s) but with no parents present. Variables include "number of children at home" by various age groups of children, and "average number of children per census family."

- c) The 2006 census distribution is used and there is a reconciliation done to ensure agreement with the previously projected count of children in the family data.

21) HOUSEHOLDS BY FAMILY TYPE

- a) The universe for "households by family type" is "occupied private dwellings" (also known as "private households").
- b) The variables include "family households" (single family and multiple family) and "non-family households."
- c) The 2006 census percentage distribution in conjunction with 2006 census family counts was used as the starting point for each year.

22) TENURE OF DWELLING

- a) The universe for "tenure of dwelling" is occupied private dwellings.
- b) The variables include owned, rented and band housing.
- c) The 2006 census tenure distribution was used to start the series and then the annualized rate of change between the DA tenure numbers for the 2001 census and the 2006 census were used to generate the subsequent estimates and projections.

23) CANADIAN CITIZENSHIP BY AGE

- a) The universe for " **citizenship** " is Canadian citizens in occupied private dwellings (where Canadian citizens were derived from the household population).
- b) The variables are 5 year age cohorts for Canadian citizens.
- c) The 2006 census Canadian citizens by age data were used to start the series. The change 2001-2006 census for citizenship rate and immigrant rate were used to derive formulas for subsequent years.

Data File Format

For DEP 2012, each variable begins with either an "E" for Estimate or "P" for Projection. Estimates are provided for the year 2012 and for 5 years earlier, 2006. Projections are provided for 2015, 2017 and 2022.

So, "E07TOTPOP" and "E12TOTPOP" are Estimated 2007 Total Population and 2012 Total Population. The abbreviations "P15TOTPOP," "P17TOTPOP" and "P22TOTPOP" are for projected total populations for 2015, 2017 and 2022.

Roll-up tables are provided for all standard geographies. The following geographies are provided:

CAN	Country
CCS	Census Consolidated Subdivision
CMACA	Census Metropolitan Area
CMACT	Census Tract
ER	Economic Region
FED03	Federal Electoral District
FSA	Forward Sortation Area
PR	Province
PRCD	Census Division
PRCDCSD	Census Subdivision
PRCDDA	Dissemination Area
SAC	Statistical Area Classification

Data File Structure

For each level of geographic detail, there are tables for the year of data in question. Due to constraints imposed by most database software, each geographic level for each year has three separate tables. The third table is new to the DEP 2012 product and was introduced to hold the new Citizenship variables without impacting the historical file structure comparability. The first field in every table is the appropriate standard geographic code. The naming convention is as follows:

DEPxx_yr_geog_p# xx = version of DEP product
 yr = year of estimate or projection
 geog = geographic level (see chart)
 p# = part identifier (1, 2 or 3)

Variables

The list of variables is available as an Excel file titled "DEP 2012 - Variables.xls".

Legal

Selected Environics Analytics Information Products are based, in whole or in part, on Computer File(s) licensed from Statistics Canada © Copyright, HER MAJESTY THE QUEEN IN RIGHT OF CANADA, as represented by the Minister of Industry, Statistics Canada 2012.

Environics Analytics Group is an Authorized User of selected Statistics Canada Computer File(s) and Distributor of derived Information Products under Licensing Agreement 6894.

Environics Analytics Group is an Authorized Reseller of selected Statistics Canada Computer File(s) under Licensing Agreement 6894.

No confidential information about an individual, family, household, organisation or business has been obtained from Statistics Canada.

Trademarks and logos of various products and companies are the property of the respective parties. PRIZM_{C2} is a product of Environics Analytics Group Ltd. The Nielsen Company is a Sales Agent of Environics Analytics in the United States.

PRIZM is a registered trademark of The Nielsen Company and is used with permission. Selected PRIZM_{C2} nicknames are trademarks of The Nielsen Company and are used with permission. Other PRIZM_{C2} nicknames are trademarks of Environics Analytics Group Ltd.