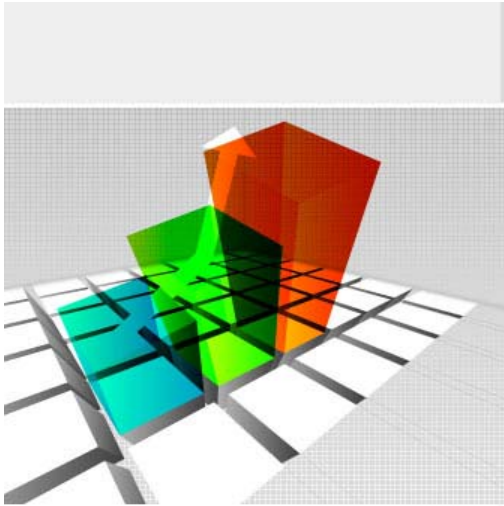




**Everything you always wanted to
know about the DHS, but you
never dared to ask**

Simona Bignami, PhD
Département de démographie
Université de Montréal

October 25, 2012

Outline

- The surveys
- The data
- The analysis
- Implications for other surveys

Outline

➤ **The surveys**

- Brief history of the DHS program
- The original goal
- How the DHS' objectives have evolved over time
- How many different types of surveys today?

➤ The data

➤ The analysis

➤ Implications for other surveys



The surveys

Brief history of the DHS program

- Begun in 1984
- Successor of World Fertility Survey and Contraceptive Prevalence Survey
- Since 1984, the DHS project has provided **technical assistance** to more than 300 surveys in over 90 countries, advancing global understanding of health and **population trends** in developing countries

The DHS' conceptual framework and original goal

- WFS, CPS: demographic surveys grounded in **Bongaarts' proximate determinants of fertility framework (1978; 1984)**
 - The tool most widely used by demographers for analyzing fertility and fertility change
 - The proximate determinants of fertility (PDF) are the biological and behavioral factors through which socioeconomic and environmental variables operate to influence the rate of childbearing in a population
 - Seven PDF that have a direct influence on fertility at the societal level:
 1. Patterns of union formation
 2. Contraceptive use
 3. Induced abortion
 4. Permanent sterility
 5. Intrauterine mortality
 6. Frequency of sexual intercourse (→ probability of conception)
 7. Lactational amenorrhea and post-partum abstinence (→ duration of post-partum infecundability)

The DHS' conceptual framework and original goal (*cont.*)

- DHS adopted the same conceptual framework as the WFS and the CPS, but also added important questions on maternal and child health and nutrition.
- Even today, the DHS reports present fertility-related information in three main sections:
 - **Fertility** (incl. age at first birth and birth intervals)
 - **Family planning** (incl. contraceptive use and sterilisation)
 - **Other proximate determinants of fertility** (incl. age at first marriage, sexual behaviour, post-partum amenorrhea, abstinence and insusceptibility)
- Key implications for data use and analysis:
 - Not possible to study marital mobility
 - Not possible to study sexual behaviour within marriage
 - Ever-married women samples at the beginning; men survey not always included (especially in earlier surveys)

How the DHS' objectives have evolved over time

- Through the DHS' first phase, most emphasis remained on **D**emography (in line with the WFS)
- More recently, and especially since the introduction of biomarkers, **H**ealth issues and access have taken a prominent role

Chronology: DHS Phase I

1984-1997 The Demographic and Health Surveys project is established at the Institute for Resource Development (IRD)

1985 Fieldwork begins in El Salvador, DHS's first survey

1987 Anthropometry is measured for first time, Sri Lanka

1989 IRD is acquired by Macro International Inc

1995 Anemia is measured for first time, Kazakhstan

Chronology: DHS Phase II

1997-2003 DHS is folded into USAID's multi-project MEASURE program as MEASURE DHS+, which incorporates traditional DHS features, expands the content on maternal and child health, and adds biomarker testing to numerous surveys

1999 First **Service Provision Assessment Survey**, Bangladesh

2001 First **HIV testing** in a DHS, Mali and Zambia

Chronology: DHS Phase III

2003-2008 In October 2003, under the new MEASURE DHS project, ORC Macro joins forces with Johns Hopkins University Bloomberg School of Public Health/Center for Communication Programs, PATH, Casals and Associates, and Jorge Scientific Corporation (JSC) to expand data collection efforts and access to and use of demographic and health data.

2006 First **Malaria Indicator Survey**, Angola

2007 PDAs are used to collect data, Peru

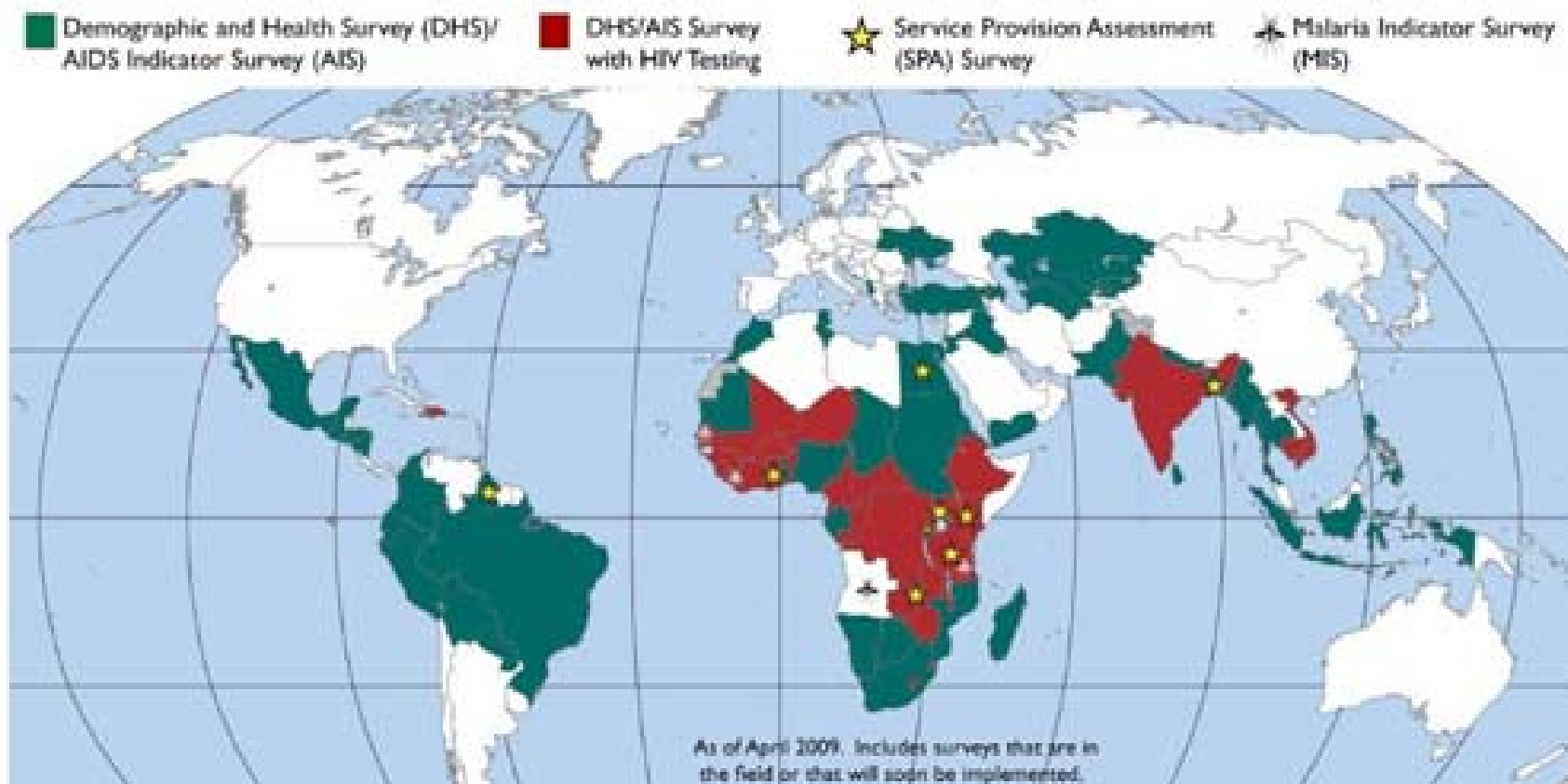
2009-2014 USAID renews funding to Macro, now ICF International

2010 DHS Celebrates 25 years

How many different types of surveys today?

- **Demographic and Health Survey (DHS):** household survey that provides data on population, health, and nutrition (including anthropometric measurements for children) and that may include testing for anemia, HIV, and/or other STIs
- **AIDS Indicator Survey (AIS):** household survey focused on HIV/AIDS that may include testing for HIV, syphilis, and/or herpes
- **Malaria Indicator Survey (MIS):** household survey focused on malaria that may include testing for malaria and/or anemia
- **Service Provision Assessment Survey (SPA):** facility-based survey that obtains information on health services including infrastructure, service availability, and quality

How many different types of surveys today?



Outline

➤ The surveys

➤ **The data**

- The questionnaires
- The DHS core questionnaires
- The DHS modules
- Special questions and topics
- How about data quality?
- The biomarker data
- How to get access to DHS data

➤ The analysis

➤ Implications for other surveys



Data

MEASURE DHS has collected, analyzed and disseminated accurate and representative data on population, health, HIV and nutrition through more than 260 surveys in over 90 countries.

The questionnaires

- Basic approach: to collect data that are comparable across countries
- To achieve this, standard model questionnaires have been developed, and form the basis for the questionnaires that are applied in each country
- **Standard model questionnaires = core questionnaire + modules**
- Countries are asked to adopt the model questionnaire in its entirety, but can add questions of particular interest and delete questions if they are considered irrelevant for their specific context

The DHS core questionnaires

- **Household questionnaire:** administered to the head of the household, collects information on all household members and is used to draw the sample for the individual interviews with women and men
- **Women questionnaire:** administered to sampled women age 15-49
- **Men questionnaire:** administered to sample men age 15-59 (or 15-54)

The DHS household questionnaire

- **Household listing:** for every usual member of the household and visitor, information is collected about age, sex, relationship to the head of the household, education, and parental survivorship and residence
- **Household characteristics:** questions ask about the source of drinking water, toilet facilities, cooking fuel, and assets of the household; in areas with a high prevalence of malaria, questions about the use of bed nets in the household are added
- **Nutritional status and anemia:** the height and weight of women age 15–49 and young children are measured to assess nutritional status; or the same individuals, the level of hemoglobin in the blood is measured to assess the level of anemia.

The DHS women questionnaire

- **Background characteristics:** questions on age, marital status, education, employment, and place of residence provide information on characteristics likely to influence demographic and health behavior
- **Reproductive behavior and intentions:** questions cover dates and survival status of all births, pregnancies that did not end in a live birth, current pregnancy status, fertility preferences, and future childbearing intentions of each woman
 - All DHS have a full birth history, which includes all live births, and a pregnancy history for the 5 years before the survey
 - A full pregnancy history is included only in a few DHS, mainly for Eastern European countries

The DHS women questionnaire (cont.)

- **Contraception:** questions cover knowledge and use of specific contraceptive methods, source of contraceptive methods, exposure to family planning messages, informed choice, and unmet needs for family planning; for women not using contraception, questions are included on knowledge of a source of contraception and intentions about future use.
- **Antenatal, delivery, and postpartum care:** the questionnaire collects information on antenatal and postpartum care, place of delivery, who attended the delivery, birth weight, and the nature of complications during pregnancy for recent births
- **Breastfeeding and nutrition:** questions cover feeding practices, the length of breastfeeding, and children's consumption of liquids and solid food
- **Children's health:** questions examine immunization coverage, vitamin A supplementation, recent occurrences of diarrhea, fever, and cough for young children and treatment of childhood diseases
- **Status of women:** the questionnaire asks about various aspects of women's empowerment, including decision making and autonomy, and about attitudes towards domestic violence
- **AIDS and other sexually transmitted infections:** questions assess women's knowledge of AIDS and other sexually transmitted infections, the sources of their knowledge about AIDS, knowledge about ways to avoid getting AIDS, and high-risk sexual behavior
- **Husband's background:** currently married women are asked about the age, education, and occupation of their husbands

The DHS men's questionnaire

- **Background characteristics:** questions on age, education, employment status, religion, and place of residence are included to provide information on characteristics likely to influence men's behavior
- **Reproduction:** data are collected on the number of children that the man has fathered in his lifetime, survival status of births, number of women he has fathered children with, antenatal and delivery care for the last child born in the previous 3 years, and man's knowledge on feeding practices for children with diarrhea; questions are also asked about fertility preferences and future childbearing intentions of each man
- **Knowledge and use of contraception:** questions are designed to determine knowledge and use of specific family planning methods; men are also asked about their exposure to family planning messages through both the media and health professionals, about the most fertile days in a woman's cycle, and condom (male and female) sources
- **Employment and gender roles:** men are asked about their employment and occupation, as well as about their attitude towards various aspects of women's empowerment, such as decision making, childbearing, women's autonomy, and domestic violence
- **AIDS and other sexually transmitted infections:** questions assess men's knowledge of AIDS and other sexually transmitted infections, the sources of their knowledge about AIDS, knowledge about ways to avoid getting AIDS, and high-risk sexual behavior
- **Other health issues:** men are asked about various health issues such as tuberculosis, circumcision, injections, use of tobacco, and health and care for their children

The DHS modules

- Developed to address the special needs of some countries on specific topics not contained in the core questionnaires, while achieving a certain level of comparability
- **Current modules:**
 - Domestic Violence
 - Female Genital Cutting
 - Maternal Mortality
 - Fistula
- Modules from previous DHS phases:
 - Consanguinity, Women's Status, Sterilization Experience, Pill failure and behavior, Malaria
- How to find out which surveys include certain modules?
 - Check online: "Survey search" ➔ "Select survey characteristics"

Special questions and topics

- Special DHS: DHS on specific topics
- Special questions: causes of death and verbal autopsies, avian influenza knowledge (Egypt 2008), health expenditures
- Special questions have been included in order to construct indicators to monitor progress towards the MDG
 - Birth registration
 - OVCs
- Special sections: reproductive calendar
- GPS data (140 surveys)

How about data quality?

- DHS data are deemed of **high quality**, but have limitations
- Most important **limitations**:
 - Cross-sectional data (exception:
 - Behaviors are generally self-reported
 - Nationally-representative (sometimes representative at the regional level as well), but not representative of certain subgroups that maybe of interest
 - Prone to known errors
 - Recall (especially of dates and ages)
 - Displacement of events in the birth history (see Machiyama, 2010)

The biomarker data

- **Anthropometrics and anemia**
- **HIV**
- **Malaria**
- **Special biomarker data collection**
 - Blood pressure (13 countries)
 - Diabetes (2 countries)
 - Hepatitis B & C (3 countries)
 - Lead testing (2 countries)
 - Lipids testing (1 country)
 - Measles (2 countries)
 - Tetanus (2 countries)

How to get access to DHS data

- Register on the website of Measure DHS
- Request datasets among those available
 - A separate request is needed for biomarker and GPS data
- Wait 24-48 hours
- Full step-by-step instructions:

<http://www.measuredhs.com/data/Access-Instructions.cfm>

Outline

- The surveys
- The data
- **The analysis**
 - Why are there more datasets than questionnaires?
 - Get to know your variables
 - Descriptive statistics always come first
 - Tips for the appropriate use of survey weights
 - The online tools
- Implications for other surveys



The analysis

Why are there more data files than questionnaires?

Data from the household questionnaire:

1. **Household Data - Household Recode (HR)** L SEP *This dataset has one record for each household.* It includes household member's roster but no information from the individual women/men questionnaires is present in this file. The unit of analysis (case) in this file is the household.
2. **Household Listing Data - Household Member Recode (PR)** L SEP *This dataset has one record for every household member.* It includes variables like sex, age, education, orphanhood, height and weight measurement, hemoglobin, etc. It also includes the characteristics of the households where the individual lives or was visiting. The unit of analysis (case) in this file is the household member.

Why are there more data files than questionnaires? (*cont.*)

Data from the individual questionnaires:

1. **Individual Women's Data - Individual Recode (IR)**  ***This***

Why are there more data files than questionnaires? (cont.)

2. **Men's Data - Male Recode (MR)** SEP *This dataset has one record for every eligible man as defined by the household schedule.* It contains all the data collected in the men's questionnaire plus some variables from the household. The unit of analysis (case) in this file is the man.
3. **Couple's Data - Couple's Recode (CR)** SEP *This dataset has one record for every couple.* It contains data for married or living together men and woman who both declared to be married (living together) to each other and with completed individual interviews (questionnaires). Essentially the file is the result of linking the two files previously described based on whom they both declared as partners. The unit of analysis (case) in this file is the couple in which both partners were interviewed.

Why are there more data files than questionnaires? (cont.)

4. **Children's Data - Children's Recode (KR)** L SEP *This dataset has one record for every child of interviewed women, born in the five years preceding the survey.* It contains the information related to the child's pregnancy and postnatal care and immunization and health. The data for the mother of each of these children is included. This file is used to look at child health indicators such as immunization coverage, vitamin A supplementation, and recent occurrences of diarrhea, fever, and cough for young children and treatment of childhood diseases. The unit of analysis (case) in this file is the children of women born in the last 5 years (0-59 months).
5. **Births' data - Birth's Recode (BR)** L SEP *This dataset has one record for every child ever born to interviewed women.* Essentially, it is the full birth history of all women interviewed including its information on pregnancy and postnatal care as well as immunization and health for children born in the last 5 years. Data for the mother of each of these children is also included. This file can be used to calculate health indicators as well as fertility and mortality rates. The unit of analysis (case) in this file is the children ever born of eligible women.

Why are there more data files than questionnaires? (cont.)

Associated files:

1. **Wealth Index data (WI)**^{LSEP} *This dataset has one record for every household.*
2. **Height and Weight data (HW)**^{LSEP} *This dataset has one record for every child measured for height and weight.*

Why are there more data files than questionnaires? (cont.)

Biomarker data:

1. **HIV Test data - AIDS Recode (AR)**^{L SEP} *This dataset has one record for every individual for which blood was drawn for HIV testing.*
2. **Other Biomarkers data (OB)**^{L SEP} *This dataset has one record for every individual for which samples were taken for different kinds of biomarkers.*

Both files can be linked to the household members (PR), the women (IR), and the men files (MR).

Why are there more data files than questionnaires? (cont.)

GPS data:

- **Geographic data (GE)** SEP *This dataset has one record for every cluster in which the survey was conducted.*
This type of file includes the latitude and longitude of the center of the sample cluster.

Get to know your variables

- **DHS recode manual = DHS analysis bible**

- **Example:**

In your dataset (assuming you are using an IR, BR, KR, or MR file) check the label of v107 (mv107). The label says "highest year of education." If you analyze this variable assuming it is the respondent's highest year of education, you will have highly misleading results. Why? Because the variable label needs to be short, and so cannot give complete information about every variable included in the dataset. Download the DHS recode manual and look through it to find v107. See that v107 is the highest year of education at the level recorded in v106. Had you analyzed v107 as the highest years of education, you would have seriously underestimated the level of education in the country you are studying.

- **Download at:**

<http://www.measuredhs.com/publications/publication-dhsg4-dhs-questionnaires-and-manuals.cfm>

Descriptive statistics always come first

- DHS Final Reports provide a wealth of descriptive statistics about the most commonly used indicators; they also provide sample sizes (denominators) for calculating them
- **You always need to make sure that your sample sizes (and indicators, if same) match those in the final report!!!**

Tips for the appropriate use of DHS survey weights

- Survey weights are used to make sample data representative of the entire population.
- **Survey weights need to be used for all descriptive and multivariate analyses**
- **Pay attention to clusters and strata!**

Unit of analysis	Weight
Households and household members	hv005
Women	v005
Men	mv005
Couples	mv005
HIV test results	
Women	hiv05
Men	hiv05
Couples	mv005

Tips for the appropriate use of DHS survey weights (*cont.*)

➤ If in doubt, refer to the Guide to DHS Statistics:

<http://www.measuredhs.com/publications/publication-dhsg1-dhs-questionnaires-and-manuals.cfm>

The online tools

- Do not underestimate the online tools provided by Measure DHS
 - Useful for a quick exploratory analysis, especially if many countries are involved
 - Useful for displaying or checking standard indicators (TFR, HIV prevalence etc.), especially if they are difficult to reproduce (TFR); full list of indicators available:
<http://www.measuredhs.com/data/Survey-Indicators.cfm>
- Do not overexploit these tools!
 - If you are dealing with subpopulations, you always need to analyze the microdata

Outline

- The surveys
- The data
- The analysis
- **Implications for other surveys**
 1. Get familiar with your survey!
 2. Understand the general data structure
 3. Always begin with descriptive statistics, and try to reproduce published figures when possible
 4. Apply appropriate weighting procedures

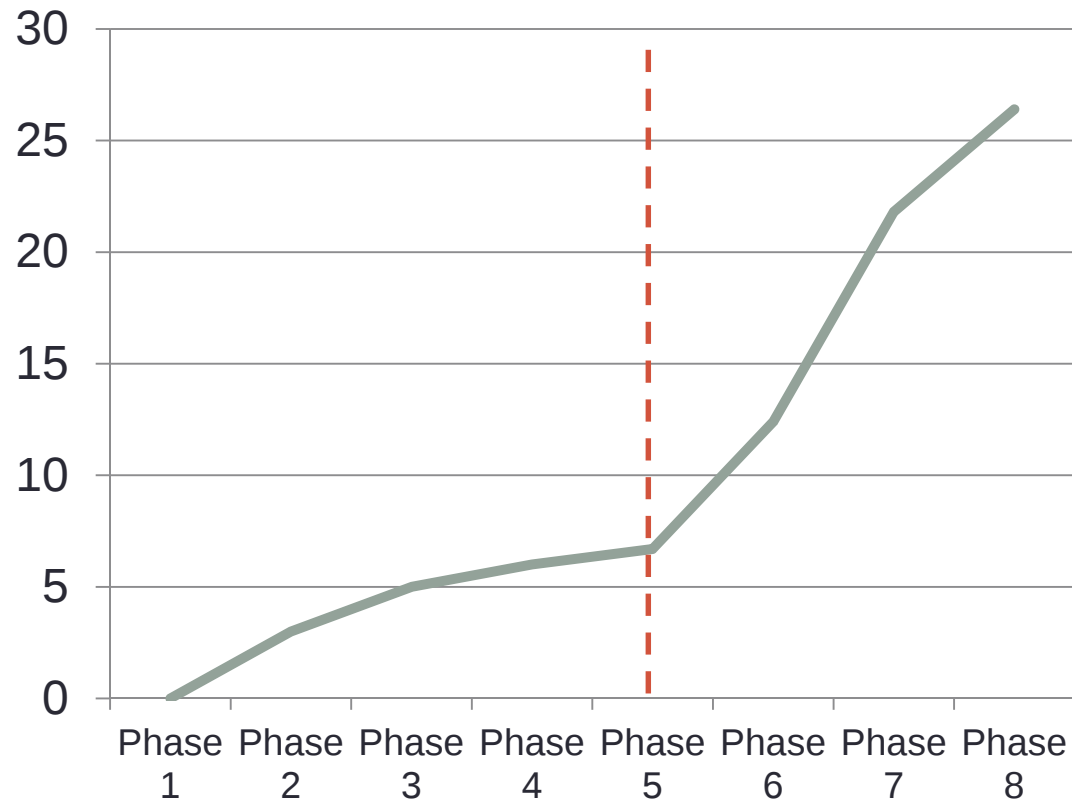


Implications for other surveys

1. Get familiar with your survey!

- Example # 1: ELDEQ
- Example # 2: Canadian Census 2011
- How many questionnaires?

**Cumulative attrition rate,
ELDEQ**



2. Understand the general data structure

- How many data files?
- Data files for the whole sample or for a sub-sample?
- Which file formats?
- What merging procedures?
 - Particularly important for longitudinal data

3. Always begin with descriptive statistics

- Maybe more difficult with certain surveys where there are less published figures
- Make sure at least to match **sample sizes**

4. Apply appropriate weighting procedures

- **All** survey data need to be analyzed by applying the most appropriate weighting procedures
 - Example: Statistics Canada's cross-sectional vs longitudinal weights