The 2008 Revision of the United Nation *World Population Prospects*: Challenges in Estimating and Projecting the World's Population.

Gerhard K. Heilig[(1)], Patrick Gerland[(2)], Nan Li[(2)],
Vladimira Kantorova[(2)], Taeke Gjaltema[(2)], Pablo Lattes[(2)],*

Short Abstract:

The United Nations Population Division has prepared global population estimates and projections since 1951. The 2008 Revision of the *World Population Prospects* (WPP) provides a broad range of demographic indicators for all countries of the world, all development groups, and major geographical regions - covering the period from 1950 to 2050. This paper outlines the basic steps in preparing the population estimates for the *World Population Prospects*. There are still a significant number of countries where demographic information is incomplete, outdated, seriously distorted or almost completely lacking. The paper describes some of the challenges in deriving consistent and plausible demographic estimates for countries with highly deficient or unreliable empirical data. It also provides an overview of demographic data availability and data quality for all countries of the world for the period since 1950.

---

(1) Chief, Population Estimates and Projections Section, United Nations Population Division

(2) Population Affairs Officers, Population Estimates and Projections Section, United Nations Population Division

Introduction

With each revision of the *World Population Prospects* (WPP) the United Nations Population Division is carrying out a complete re-estimation of historical demographic trends since 1950 for all countries and territories of the world. These population estimates are based on the most recently available data sources, such as censuses, demographic surveys, data from vital and population registers and various other sources. With each new data collection the time series of fertility, mortality and migration, as well as population trends by age and sex can be extended and, if necessary, retrospectively corrected. For countries with highly deficient demographic data or countries without recent census, newly conducted demographic surveys can often help to reconstruct historical demographic trends much more accurately than was possible in previous revisions. In the 2008 revision of the WPP, this re-estimation of population by age and sex, fertility, mortality, and migration covers the period between 1950 and 2008 – which is the base-year of the projections. The following paper describes the major steps in this estimation process and highlights some of the challenges.

Each WPP re-estimation process consists of four major steps:

- **Data collection**: Country analysts systematically collect all newly available data from censuses, surveys, vital and population registers, analytical reports and other sources. Traditionally, the data are provided by the UN Statistics Division (Demographic Yearbook), by the Regional Commissions, by other UN agencies (WHO, UNICEF, UNAIDS) and by international data archives (DHS). Increasingly, analysts also search for demographic information on the web sites of national statistical offices and contact colleagues in these offices to send CD-ROMs and statistical reports. Ideally, each new data set for a country would include a census population by (single years of) age and sex, a recent life table, age-specific fertility rates, and net-migration by age and sex.
- **Checking of completeness and demographic plausibility**: In a second step these data are evaluated for geographical completeness and demographic plausibility. Post-enumeration surveys are used to evaluate census data. If necessary, adjusted data are acquired or adjustments are applied using standard demographic techniques, such as for under-enumeration of young children or age-heaping. For countries were no, or only minimal, demographic information is available demographic models are used to estimate fertility, mortality and migration. Previous estimates are reviewed and, if necessary, revised according to the new data.
- **Checking consistency between demographic components**: These new provisional country estimates are then subjected to a series of checks in which the relationship between the demographic components (fertility, mortality, migration) is tested. For a growing number of countries, where several censuses since the 1950s are available, inter-censal consistency is analyzed by projecting the population between census years – using the provisional estimates for fertility, mortality and migration. Essentially, this means that the base-year of the projection is pushed backward to the 1970s, 1960s or even 1950s. If the population by age and sex of the subsequent censuses cannot be matched by the projection, adjustments for one or more demographic components are made. In some cases the initial census population itself may be revised after back-surviving cohorts from one or multiple subsequent censuses with more reliable results.

- **Checking consistency between countries**: Finally, all countries are projected and the results aggregated to the required geographical and economic regions. In a final round of checking, the estimation and projection results of all countries of a particular region and at a particular initial level of fertility and mortality are compared to each other. "Outliers" and "crossovers" in estimated and projected trends are identified and, if necessary, adjusted. This should help to avoid that countries with similar level of fertility and mortality, or countries within the same geographic region, have very divergent estimates or projections without clear justification. An important aspect of these final adjustments is the consolidation of international migration, which has to be zero for the world.

In this process of demographic estimation analysts typically are facing several methodological challenges which are discussed below:

1. Data availability

Despite significant efforts in the development of statistical systems, demographic information is still highly deficient or lacking for large parts of the (developing) world. Roughly speaking, for some 60% of the countries of the world, and for one-third of the world's population, the available demographic information is still deficient in some respect and has to be evaluated, cross-checked, complemented with information derived from demographic models or adjusted according to standard demographic techniques. For some of these countries, the evaluation and estimation process may be relatively straight forward; however for about 30% of the countries of the world demographic data are so deficient, inconsistent or even lacking that major re-estimation procedures are necessary to derive the input data that are necessary for carrying out projections. With *reported* demographic statistics alone, cohort-component population projections could not be carried out for a *majority* of the countries of the world.

Table 1 and 2 provide a synopsis of demographic data availability for the 2008 Revision of the *World Population Prospects*. These tables were derived from internal meta-information of data sources by demographic component (for details of the method see Footnote of Table 1). Based on our experience, roughly 14% of the countries (which represent 4.5% of the world population) have very poor or lacking demographic data. In these countries, data for at least one, but typically more than one, demographic component are older than 15 years, highly deficient or completely missing. Another 11% of the countries (with an additional 6% of the world population) have *some* data for each demographic component, but these data are 10 to 15 years old. In addition, more specific information, such as age-specific fertility rates are missing or are older than 10 years. For 35% of the countries of the world (representing 23% of the world population) we have basic demographic data for each component; however, this information typically consists of highly aggregated summary statistics or absolute numbers, such as total fertility, total number of birth or total number of deaths. *Age-specific* rates, or complete life tables are not available for these countries (see table 1 and 2).

Population censuses, which are corner stones for the estimation of demographic trends, have not been carried out in a significant number of countries in recent decades. Based on an inventory of the UN Statistics Division in 2008, 25 countries (of 224 countries reviewed) have not carried out censuses in the 2000 census round (between 1995 and 2004) – 16 of them in Africa and 6 in Asia (see also figure 1). Hopefully, the 2010 census round of censuses will improve this situation.

These numbers on data availability clearly indicate that, rather than just collecting readily available demographic information, significant analytical efforts are necessary for the preparation of the Population Division's world population estimates and projections.

2. Data coverage and quality

Even those countries, which have carried out population censuses in recent years, rarely provide information about the quality of these data. The results of post-enumeration surveys and consistency checks are hardly ever published or accessible from outside the country. Often it is unclear whether, and to what extent, census data have been adjusted and if so, what adjustment methodology was used.

A number of countries can not provide demographic data for their whole territory due to logistic problems, political conditions or violent conflicts. A similar challenge is the acquisition of demographic information for disputed territories. These kinds of problems have to be confronted in countries such as Morocco (Western Sahara), Sri Lanka, Serbia, Montenegro, Moldova, Cyprus, Israel (Status of Jerusalem), Sudan (South Sudan, Darfur), Azerbaijan, or Georgia.

Sometimes census results are not, or only partially, published due to political reasons or lack of resources. Quite often, countries have only released highly aggregated statistics, such as total population by province. Apparently, population by (single years of) age and sex is sometimes not published, because the data would reveal unexplainable flaws in the census, such as underreporting of young children, age heaping, or serious imbalances in the sex ratios at birth.

Some countries have a history of disputed censuses: Results of the first post-independence census of Nigeria conducted in 1962 were withdrawn due to serious doubts about its validity and the 1973 census was never published because its results were discredited. The 1991 census, which put Nigeria's population at 88.5 million, was also disputed and even Nigeria's most recent census in 2006 has been mired in controversy and allegations of fraud. In these highly controversial and politically charged cases it is extremely difficult for the Population Division to derive impartial and reliable demographic estimates. Similar challenges consists in countries such as Iraq, Afghanistan, or Sudan.

3. Consistency between data sources

One of the most serious challenges in estimating the world's population is the inconsistency in demographic data from different sources. It is not uncommon in many (developing) nations that different official institutions, such as the central bureau of statistics, the statistical offices of ministries, government planning boards or state-run family planning institutions are publishing widely diverging estimates of fertility, mortality or population by age and sex. Moreover, in recent decades a growing number of countries have participated in various international programs of demographic data collection, such as in the USAID-funded Demographic and Health Surveys (DHS), the Multiple Indicator Cluster Surveys (MICS) from UNICEF, the Pan Arab Project for Child Development and the Pan Arab Project for Family Health (PAPCHILD / PAPFAM) from the Arab League, the international Reproductive Health Surveys (RHS) from the US Center of Disease Control (CDC) and the World Health Survey from WHO. In addition, a whole range of national demographic and social surveys is available. It is a major challenge just to keep track of the latest available rounds of the various surveys. In the DHS program alone, more than 200 surveys in 75

countries have been completed. Even more challenging is the fact that results from these surveys often lead to rather different estimates of mortality and fertility *for the same time periods*, even if the same methodology was used for estimation (children ever born, children surviving, etc.). Typically the various data sources plotted together in one chart produce "clouds of data points", rather than nicely aligned time series of demographic indicators (see figure 2).

4. Estimation of mortality

A core element of the WPP preparation process is the estimation of mortality trends. Ideally, complete life tables should be used. Unfortunately, reliable life tables for a significant number of years in the estimation period are only available for a minority of the countries of the world.

The most consistent international collection of life tables is included in the Human Mortality Database (HMD) created and maintained by the Department of Demography at the University of California, Berkeley (USA) and the Max Planck Institute for Demographic Research in Rostock (Germany). As of mid-2008 it provided data for the following 34 countries: Australia, Austria, Belarus, Belgium, Bulgaria, Canada, Czech Republic, Denmark, England and Wales, Estonia, Finland, France, Germany (East and West), Hungary, Iceland, Italy, Japan, Latvia, Lithuania, Luxembourg, Netherlands, New Zealand, Norway, Poland, Portugal, Russia, Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Taiwan, Ukraine and the USA. These countries represent some 17% of the world population in 2008.

To estimate mortality for the remaining more than four-fifths of the world population is a much greater challenge – particularly for those countries in Africa and Asia, where very little or almost no mortality information is available. For some regions rudimentary vital statistics (typically only operating in urban areas or densely populated parts of these countries) might be available that allow to reconstruct life tables. However, for the vast majority of these countries, mortality is estimated on the basis of infant and child mortality data provided by various data sources – using model assumptions for the underlying age patterns. The data may include direct and indirect measures from censuses and surveys. The Population Division is collaborating since 2004 with WHO and The World Bank with an inter-agency process led by UNICEF that produces common child mortality estimates (see web site: http://www.childinfo.org/mortality_igme.html) (See also: UNICEF, WHO, The World Bank and UN Population Division, 2007).

Figure 2 displays most recent child mortality (5q0) data for Afghanistan. For the year 2000, estimates range from about 75 to more then 250 deaths per 1000. Obviously, the data sources with the highest child mortality estimates would lead to completely different estimates and projections of mortality for Afghanistan than the sources with lowest data. Various techniques are used to derive 5q0 estimates from this "clouds of points" – including the weighting of data points from different sources based on their perceived reliability with subsequent calculation of a (double) Spline function or various alternative methods, such as smoothing with Loess regression.

The "cloud" of empirical child mortality data from Afghanistan in figure 2 is, by no means, an extreme case. For the majority of countries in Africa and Asia we have collected empirical information on child mortality that differs *significantly* between sources. It is one of the major

challenges in the WPP estimation process to derive plausible child mortality estimates *by sex* from these sources. These estimates are then used in model life tables to calculate age-specific survival rates, which are needed for the projections.

5. Estimation of fertility

The difficulties in estimating age-specific fertility rates for all countries of the world are similar to the challenges of estimating mortality. While fairly good fertility data are available from vital registers for the developed regions and some emerging countries, the information is much less reliable for many developing nations. There are only a few cases, where no information on fertility is available at all, but many countries publish only highly aggregated data, such as the *total* number of births.

Even if age-specific fertility rates are available, results are often questionable or highly unlikely. China's latest officially published total fertility of 1.38 (2006 Population Change Survey) is widely considered much too low. Based on the 1990 census, total fertility was estimated at 2.24 while the Population Change Survey in 1994 resulted in total fertility estimates of 1.86. The 2000 census gave a TF of 1.22. Our estimate for the year 2000 in the WPP2006 was 1.7, respectively. Government agencies within China and academic researchers have published estimates for around 2005 that range from 1.4 to 1.9. Obviously, even slight differences in fertility estimates would result in huge differences in future population size for a country with currently over 1.3 billion people.

As is well known from the demographic literature, fertility estimates for a particular period can be distorted by timing effects. If a large number of women decide to postpone childbirth during a particular historical period, fertility rates during this period could drop significantly, even if overall cohort fertility would remain unchanged. Recently reported very low fertility rates for some developing countries, notably China, may be affected by delayed childbirth and might increase again as these women have children at higher ages to reach their desired family size. In the past decade rapid economic development in some developing countries such as China or India have triggered massive internal migration from rural to urban areas – particularly among young, single women who staffed the rapidly expanding manufacturing and service sectors. These women are typically postponing marriage and childbirth, but there is great uncertainty about their completed fertility. They may return home to their village, marry and have children – particularly in times of a global economic crisis; or they may stay to a large proportion in the cities, remain single and childless. Fertility estimates in countries with extremely rapid economic and social change are a "snapshot" at best. Only time will tell, what level of fertility this new generation of women will have at the end of their reproductive period.

6. Estimation of migration

There can be no doubt that the estimation of international migration is the "weakest link in the chain" of challenges for deriving reliable population estimates. There are both conceptual as well as technical difficulties in the preparation of reliable migration estimates; and these difficulties are, by no means, restricted to developing countries. Typically, international migration is defined as a relocation of place of residence across international borders for more than 12 months. However, international labor migrants may well stay for two or three years in a foreign country and then return home to their families on a regular basis (circular

migration). Should they be counted in their homeland, where they may actually have their center of their live, including their families? Or should they be counted as part of the population in their temporary host country. Following the de-facto concept, the Population Division attempts to estimate population *including* (semi-) permanent labor migrants – but several countries, particularly those with very high percentage of foreign labor migrants, such as Qatar, prefer population estimates based on a de-jure definition, which would exclude the labor migrants.

Conceptual problems also arise from the fact that migration statistics are fragmented into various, often incompatible segments that deal with specific migration flows, such as the UNHCR statistics of refugees and displaced persons. As a general rule, the Population Division estimates refugees based on statistics provided by UNHCR, applying the assumption that refugee populations are "returned" to their home country over a period of 10 to 15 years. However, there are several countries with ongoing wars or political conflicts where this assumption seems to be unrealistic or even impossible to be applied.

One of the most serious technical problems derives from the fact that most countries do not keep statistics of people who permanently leave the country. Immigration statistics in many European countries may be far more reliable than emigration statistics, because out-migrants typically have no incentive to report their departure to the government agencies of the former host country. Germany has recently reported that its household registration system may significantly underestimate the number of out-migrants, so that the next census may find population numbers to be lower by about 1.5 to 2 million than reported by the annual estimates from the population register.

A fundamental methodological problem in estimating migration is related to the way in which migration is statistically recorded. Typically, countries do not report (annual) flows of in- and out-migration, but only record the stock of foreign-born population or population with previous residence outside a country. So far, no one has managed to generate a reliable matrix of recent international migration flows for all (or at least the major) countries of the world – not to speak about historical time series since 1950. Even more difficult is the projection of such migration flows for the future. The greatest challenge in estimating migration flows is to consolidate incompatible migration statistics between sending and receiving countries, such as Mexico and the United States of America, or Romania and Spain. Typically, the immigration data of the receiving country do *not* match the emigration data of the sending country.

The Population Division estimates *net-migration* flows by age and sex based on all available data sources (See: Table 3). In many cases the only option available is to calculate net-migration based on the residuals between population estimates (based on fertility and mortality) and the reported population in a particular year or time period. For a significant number of countries, models (Castro-Rogers) have to be used for estimating the age-and sex-composition of net-migration due to the lack of appropriate data from countries.

7. Problems with register-based demographic data

Officially reported demographic statistics from developed countries with existing population and vital registers are often considered the "gold standard" in demographic information, and critics of the *World Population Prospects* sometimes point out divergences between UN estimates and these reported statistical data. However, these apparent discrepancies often disappear, when the National Statistical Offices *retrospectively* correct their own

demographic statistics after a census. As figures 3, 4 and 5 illustrate, there are consistent patterns of retrospective correction of population estimates in many developed countries, because register-based population statistics often under- or over-estimate migration flows. Especially in countries with large migration, register-based demographic information may be severely distorted during the inter-censal period. This can also affect fertility and mortality statistics in developed countries, due to inflated or underestimated population numbers in the denominator of fertility and mortality rates.

8. Changing challenges in data availability and estimation strategies

The preparation of the *World Population Prospects* by the United Nations Population Division has been one of the most consistent demographic research activities. For almost 60 years, the Population Division has collected and evaluated*, on a regular basis*, demographic data for all countries of the world. This gives the Division unique experience concerning *changes* in the availability and quality of international demographic statistics. Roughly two phases can be distinguished: The early period from the 1950s to the late 1970s, and the later period since the early 1980s.

- During the first period, very few countries had conducted more than one or two census and population registers, as well as vital registration systems, were not implemented in the great majority of the countries of the world. Between 1950 and 1980 65 countries had become independent and conducted their first post-colonial population census. During that phase the main methodological challenge was to evaluate the data quality of one or two demographic data collections (such as one or two censuses or surveys). Numerous methods and theories were developed to estimate and project population trends based on very limited data (United Nations, Department of Economic and Social Affairs, Population Branch, 1983).
- The situation changed fundamentally in the second phase since the early 1980s, when the results of demographic and household surveys became available, such as the results from the Demographic and Health Surveys (DHS), the Multi-Indicator Cluster Surveys (MICS) and others data collections. In addition, more countries established population and vital registers and some began to publish annual population estimates. The increasing number of data sources for one particular country now made it possible to reconstruct consistent time-series over several decades.

While in the first period population estimates were focused on period-specific analyses using one particular census, the main challenge in the second period was the analysis of the consistency between data collections for the same reference period. Cohort-analyses became much more important, as longer time-series of several censuses or surveys could be used. Since then, the challenge in population estimates for the majority of developing countries is to estimate the demographic components of mortality, fertility and migration for subsequent years in such a way, that the cohorts can be followed from one census to the next. This growth-balance method is now increasingly used by the Population Division to reconstruct complete population histories not only for developed countries (where the data are easily available), but also for a growing number of developing countries with multiple censuses and surveys (See Table 4).

As the number (and hopefully the quality) of data sources has increased in recent decades, the complexity of the estimation process has grown accordingly: Typically, inter-censal estimates of fertility, mortality and migration available from vital statistics, surveys, or other sources

cannot reproduce population change between censuses without adjustments (United Nations, Department of Technical Co-operation for Development and Statistical Office, 1982). Each new census and survey must be evaluated in the context of previous censuses and surveys. These necessary adjustments may be made to estimates of migration – often based on the residuals between projected and actual population of a particular census year. However, there may be also distortions in the reported age structure of one or more of the census populations, such as the typical underreporting of infants (Ewbank, 1981). As Ansley Cole pointed out in UN Manual IV nearly half a century ago, South Asian censuses tend to undercount infants aged 0-1, and in particular female infants (United Nations, Department of Economic and Social Affairs, Population Division, 1967), but there are many other possible distortions such as underreporting of young men, who may be migrant workers in a foreign country. Moreover, censuses in both developing and developed countries can be affected by political, organizational and methodological problems. The evaluation of census data, based on available post-enumeration surveys and methodological checks of data quality has become a major component in the preparation of population estimates for the *World Population Prospects* (See: Marks, E.S. 1978, Marks, E.S.and J.C. Rumford. 1978, Som, R.K. 1973, Statistics South Africa. 2003, Zitter, M.and E.K. McArthur. 1980, US Department of Commerce, Bureau of the Census, 1985).
Finally, the reported fertility, as well as the (infant- and child-) mortality, may also require adjustments – particularly if the reported numbers are way off the regional average, display sudden, unexplainable changes over time, or are very different between data collections. With the increasing number of data sources, population estimation has become a juggling act, where many balls are "in the air" at the same time and have to be processed simultaneously.

9. Responding to New Challenges in Population Estimation

With the increasing number of data collections from censuses, surveys, registers and other sources in recent decades from developing countries, population estimation has become more challenging. While in the past demographic *models* were often used to estimate mortality, fertility, migration – and even the age distribution of populations, actual empirical data have to be analyzed today. Contrary to popular believe this has not simplified the task of estimating consistent population trends. Just the contrary – the frequent differences between empirical data sets and the inconsistencies of data collections over time make it usually necessary to apply numerous checks and, if necessary, adjustments. This requires new tools to store and analyze the large amount of empirical demographic data. The Population Division has started a major project to develop an integrated empirical demographic database (DemoData), which will eventually hold empirical data on population by age and sex, fertility, mortality and migration for all countries of the world since 1950. This database will not only include census results, but also results from surveys and population and vital registers, as well as other sources. A first implementation of the database is now operational and was already used in the preparation of the 2008 Revision of the *World Population Prospects* (See figures 6, 7 and 8).

Conclusion

This paper highlights some of the methodological and technical challenges in the preparation of population estimates for all countries of the world for the period between 1950 and 2008. While the international development community is increasingly focusing on results-based

strategies that can be measured by internationally agreed development goals, including the Millennium Development Goals, key demographic indicators, such as infant and child mortality, life expectancy, adult mortality, and even the population of specific age groups (school children), are still highly uncertain for a significant number of (developing) countries. More than ever, there is a need for a systematic estimation and validation of demographic indicators within a consistent analytical framework that simultaneously estimates all demographic components. This estimation process is especially relevant for those 30% of the world population, primarily in developing countries, for which demographic information is still insufficient or lacking.

References:

Ewbank, Douglas C., 1981: Age Misreporting and Age-Selective Underenumeration: Sources, Patterns, and Consequences for Demographic Analysis. Washington (National Academy Press, Committee on Population and Demography, Report No. 4)

Marks, E.S. 1978. "The role of dual system estimation in census evaluation." Pp. 156-188 in Developments in dual system estimation of population size and growth, edited by K.J. Krótki. Edmonton: University of Alberta Press.

Marks, E.S.and J.C. Rumford. 1978. "The 1974 Post-Enumeration Survey of Liberia - A New Approach." Pp. 189-204 in Developments in dual system estimation of population size and growth, edited by K.J. Krótki. Edmonton: University of Alberta Press.

Som, R.K. 1973. Recall lapse in demographic enquiries. New York,: Asia Pub. House.

Statistics South Africa. 2003. "Census 2001: How the count was done." Pp. 50. Pretoria: Statistics South Africa (StatsSA).

UNICEF, WHO, The World Bank and UN Population Division, 2007: Levels and Trends of Child Mortality in 2006: Estimates developed by the Inter-agency Group for Child Mortality Estimation, New York, 2007. - http://www.childinfo.org/files/infant_child_mortality_2006.pdf and Online database http://www.childmortality.org/

United Nations, Department of Economic and Social Affairs, Population Branch, 1967: Methods of Estimating Basic Demographic Measures from Incomplete Data. Manual VI, New York

United Nations, Department of Economic and Social Affairs, Population Branch, 1983: Indirect Techniques for Demographic Estimation. Manual X, New York

United Nations. Department of Economic and Social Affairs, Population Division, 2006. *World Population Prospects*. The 2006 Revision, New York

Zitter, M.and E.K. McArthur. 1980. "Census undercount: the international experience." Pp. 164-180 in Conference on Census Undercount : proceedings of the 1980 conference. City: Washington, D.C, edited by United States. Bureau of the Census. U.S. Dept. of the Commerce, Bureau of Census: for sale by the Supt. of Docs.

US Department of Commerce, Bureau of the Census, 1985: Evaluating Censuses of Population and Housing. Statistical Training Document, ISP-TR-5

United Nations, Department of Technical Co-operation for Development and Statistical Office, 1982: National Household Survey Capability Programme. Non-sampling Errors in Household Surveys: Sources, Assessment and Control. New York

Table 1: Demographic information available for the 2008 Revision of the World Population Prospects

Number and quality of data sources	Number of countries in ...			Percentage of countries in		
	More Developed Regions	Less Developed Regions	World	More Developed Regions	Less Developed Regions	World
(1) Very poor or lacking	2	25	27	4.4	16.6	13.8
(2) Poor	1	21	22	2.2	13.9	11.2
(3) Basic	3	66	69	6.7	43.7	35.2
(4) Good	9	24	33	20.0	15.9	16.8
(5) Excellent	30	15	45	66.7	9.9	23.0
Total	45	151	196	100.0	100.0	100.0

Note: Only countries with more than 100.000 inhabitants included

Number and quality of data sources	Population in ...			Percentage population in ...		
	Developed Regions	Developing Regions	World	Developed Regions	Developing Regions	World
(1) Very poor or lacking	13,652	284,527	298,179	1.1	5.2	4.5
(2) Poor	407	393,327	393,734	0.0	7.2	5.9
(3) Basic	7,274	1,540,284	1,547,558	0.6	28.3	23.2
(4) Good	423,453	1,925,123	2,348,577	34.6	35.3	35.2
(5) Excellent	777,758	1,304,171	2,081,929	63.6	23.9	31.2

Note: Only countries with more than 100.000 inhabitants included

Note: Table 1 and table 2 are is based on the detailed meta-information about available data sources, which is generated during the estimation process for the WPP 2008 Revision in the ShortNotes for seven demographic indicators - total population, age- and sex-specific population structure, total fertility, age-specific fertility rates, total mortality, child mortality and age-specific mortality. This meta-information is evaluated according to the following criteria for the last data point available for each component:
(1) Countries, for which data on either total population, or age and sex specific population structure, or total fertility, or all mortality indicators are missing or older than 15 years. Thus, for one of the basic demographic components we have no empirical data at all or data that are outdated. Demographic models can be used to derive estimates.
(2) Countries that have some data on total population, age and sex specific population structure, total fertility and some information on mortality; but at least in one component this data are 10 to 15 years old. Detailed information (i.e. age-specific fertility rates or age-specific mortality rates) are missing or older than 10 years.
Special demographic techniques and models must be used to estimate at least one of the components or to derive for specific information necessary for the projections.
(3) Countries that have data on total population, age and sex specific population structure, total fertility and some data on mortality, none of them older than 10 years.
Detailed information, however, such as age-specific fertility rates or age-specific mortality, is missing or older than 10 years. Demographic models are often used for countries in this category to complement information available.
(4) Countries with data on all seven indicators and none are older than 10 years. Methods for extrapolation must be used to derive more recent data points.
(5) Countries with data on all seven indicators and none are older than 5 years. The empirical data can be used directly.

Table 2: Demographic information available for the WPP 2008 Revision by geographical region

	Excellent	Good	Basic	Poor	Very poor or lacking	Total	Excellent	Good	Basic	Poor	Very poor or lacking	Total
Countries	**Number**						**in percent**					
World	45	33	69	22	27	196	23.0	16.8	35.2	11.2	13.8	100
More developed regions	30	9	3	1	2	45	66.7	20.0	6.7	2.2	4.4	100
Less developing regions	15	24	66	21	25	151	9.9	15.9	**43.7**	**13.9**	**16.6**	100
Africa	2	1	26	10	16	55	3.6	1.8	**47.3**	**18.2**	**29.1**	100
Asia	5	14	21	5	5	50	10.0	28.0	**42.0**	10.0	10.0	100
Europe	26	8	3	1	2	40	65.0	20.0	7.5	2.5	5.0	100
Latin America and the Caribbean	7	7	14	5	4	37	18.9	18.9	37.8	13.5	10.8	100
Northern America	2	0	0	0	0	2	100.0	0.0	0.0	0.0	0.0	100
Oceania	3	3	5	1	0	12	25.0	25.0	41.7	8.3	0.0	100

Population	**in 1000**						**in percent**					
World	2,081,929	2,348,577	1,547,558	393,734	298,179	6,669,976	31.2	35.2	23.2	5.9	4.5	100
More developed regions	777,758	423,453	7,274	407	13,652	1,222,544	63.6	34.6	0.6	0.0	1.1	100
Less developed regions	1,304,171	1,925,123	1,540,284	393,327	284,527	5,447,432	23.9	35.3	**28.3**	**7.2**	**5.2**	100
Africa	2,068	33,858	438,865	317,056	173,032	964,880	0.2	3.5	**45.5**	**32.9**	**17.9**	100
Asia	1,193,060	1,812,054	859,286	56,593	108,671	4,029,665	29.6	45.0	21.3	1.4	2.7	100
Europe	414,134	295,487	7,274	407	13,652	730,953	56.7	40.4	1.0	0.1	1.9	100
Latin America and the Caribbean	108,801	200,411	241,012	18,839	2,823	571,886	19.0	35.0	42.1	3.3	0.5	100
Northern America	338,702	0	0	0	0	338,702	100.0	0.0	0.0	0.0	0.0	100
Oceania	25,163	6,767	1,120	839	0	33,890	74.3	20.0	3.3	2.5	0.0	100

Table 3: Data Source and Types of Migration Statistics available for the Preparation of the 2008 Revision of the *World Population Prospects*

	Type of migration data used							
	Number of countries	Official migration statistics	Migration statistics to MDC	Refugees (UHHCR)	International Net-migration Estimates	Number of different types of data used by country		
						Only one	Two	Three
Africa	55	11	8	36	16	42	10	3
Asia	50	17	4	17	36	31	14	5
Europe	40	21	0	4	33	22	18	0
Latin America	37	20	14	1	30	17	12	8
North.America	2	2	0	0	2	0	2	0
Oceania	12	3	0	0	12	9	3	0
Total	196	74	26	58	129	121	59	16

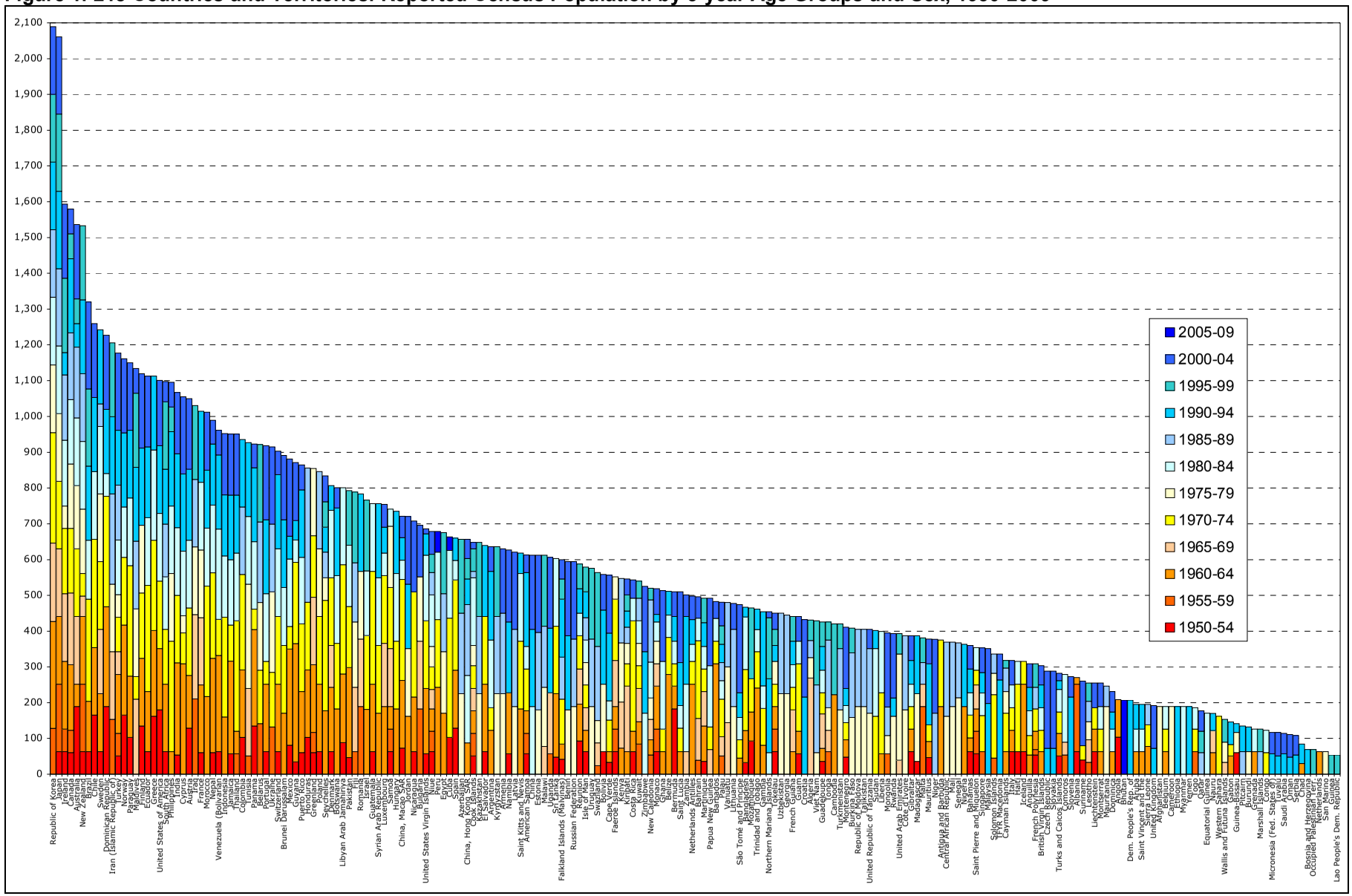
Data Availability as of mid-2008

Table 4: Distribution of countries by year of population censuses and demographic surveys (as of mid-2008)

	All countries/areas		Only countries or areas with 100,000 inhabitants or more in 2007	
	Censuses	**Demographic surveys**	**Censuses**	**Demographic surveys**
By decade				
1960-1969	256	34	182	31
1970-1979	293	134	215	127
1980-1989	262	229	192	221
1990-1999	234	434	176	427
2000-2009 (*)	253	460	197	451
By census round				
1955-1964	249	23	181	22
1965-1974	270	73	192	65
1975-1984	267	169	197	164
1985-1994	254	318	190	312
1995-2004	250	566	189	556

(*) Including planned for 2008-2009

Figure 1: 218 Countries and Territories: Reported Census Population by 5-year Age Groups and Sex, 1950-2009



Note: The figure displays the number of all data points for population by age and sex (in five year age groups) that had been reported since 1950 to the UN Statistics Division and had been published in the Demographic Yearbooks. The figure illustrates the significant differences in actual data availability from censuses between countries.

Figure 2: Child Mortality Estimates for Afghanistan (5q0) due to different data sources

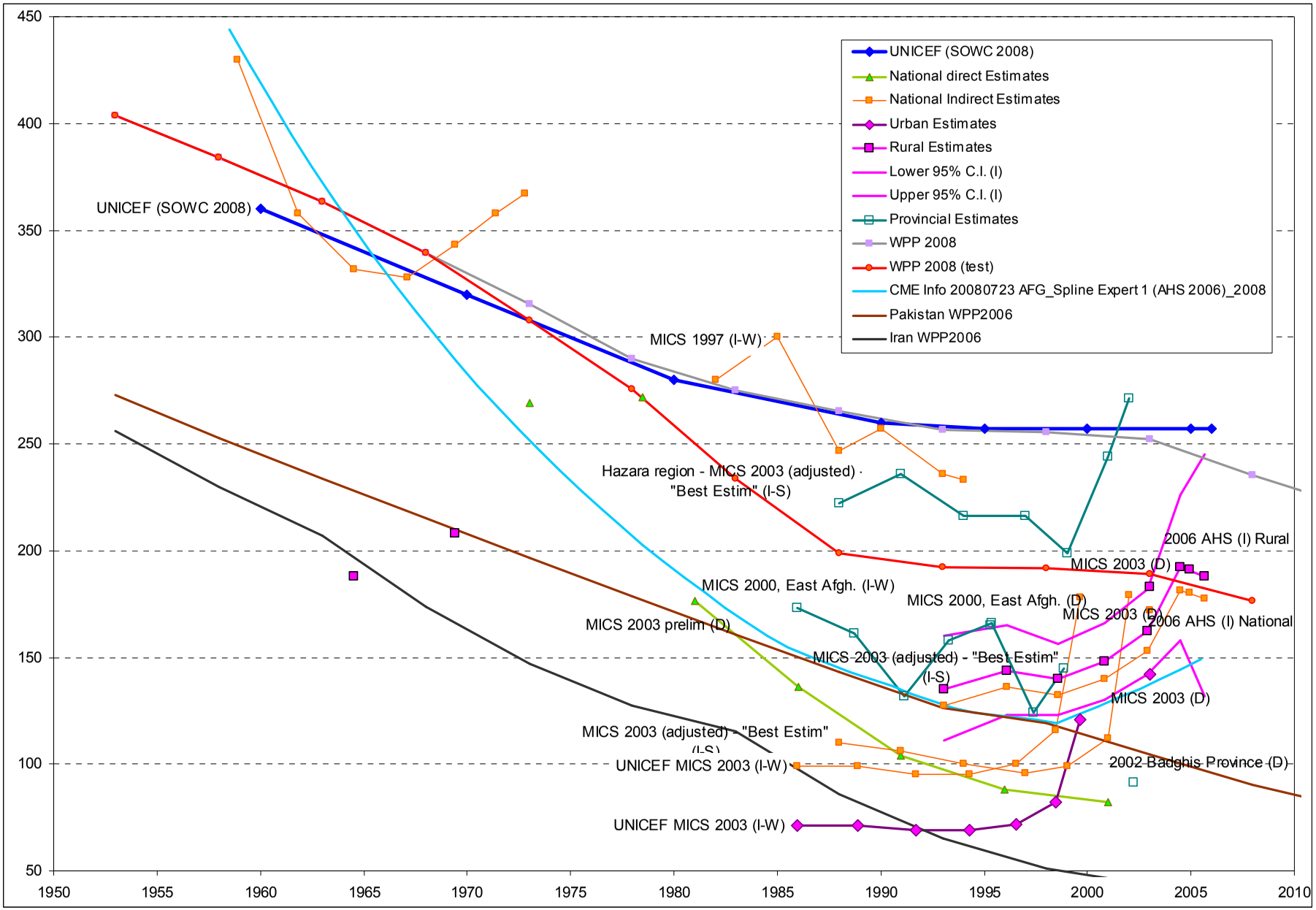


Figure 3: Barbados: Retrospective correction of population estimates after census:

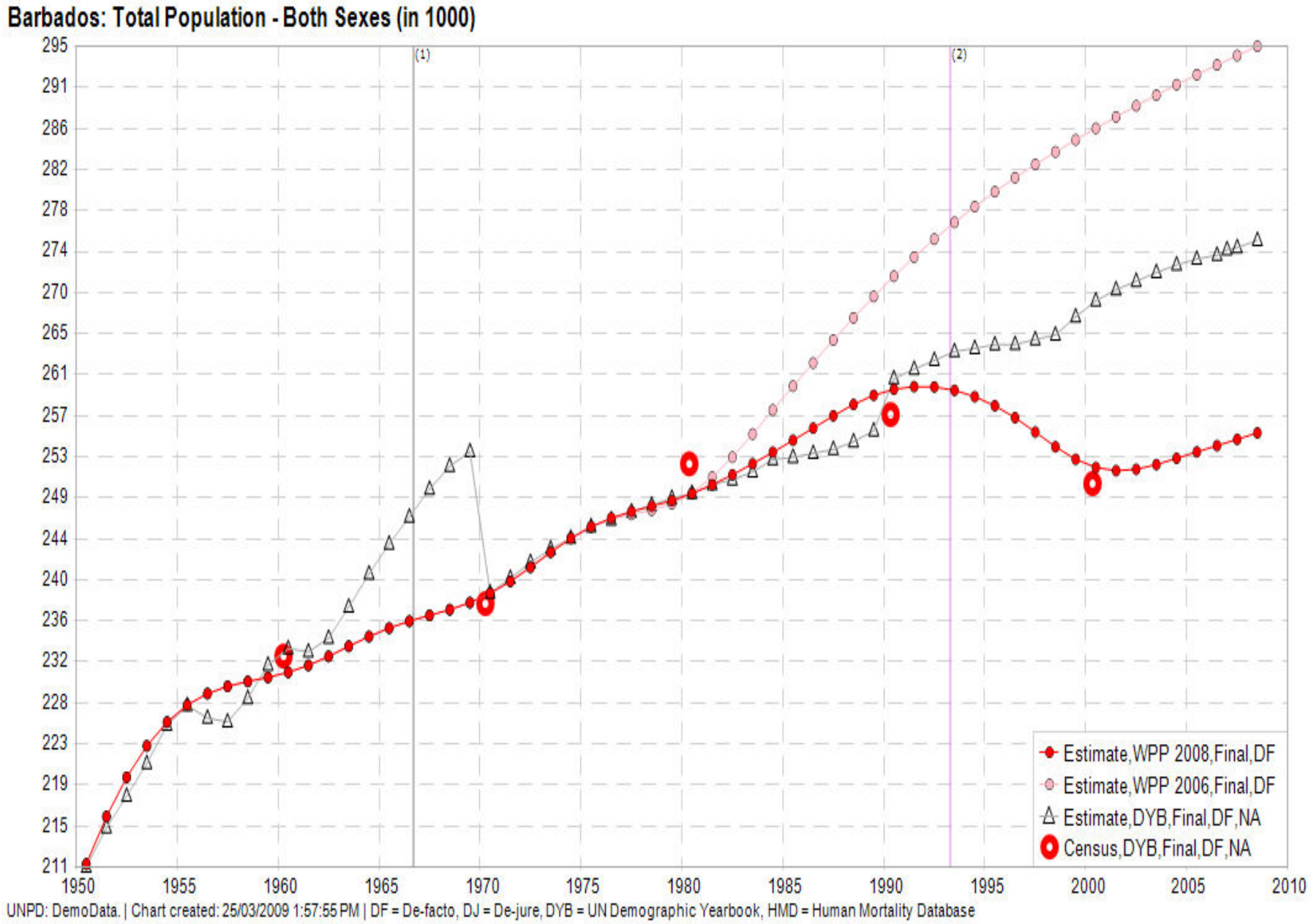


Figure 4: Czech Republic: Retrospective correction of population estimates after census:

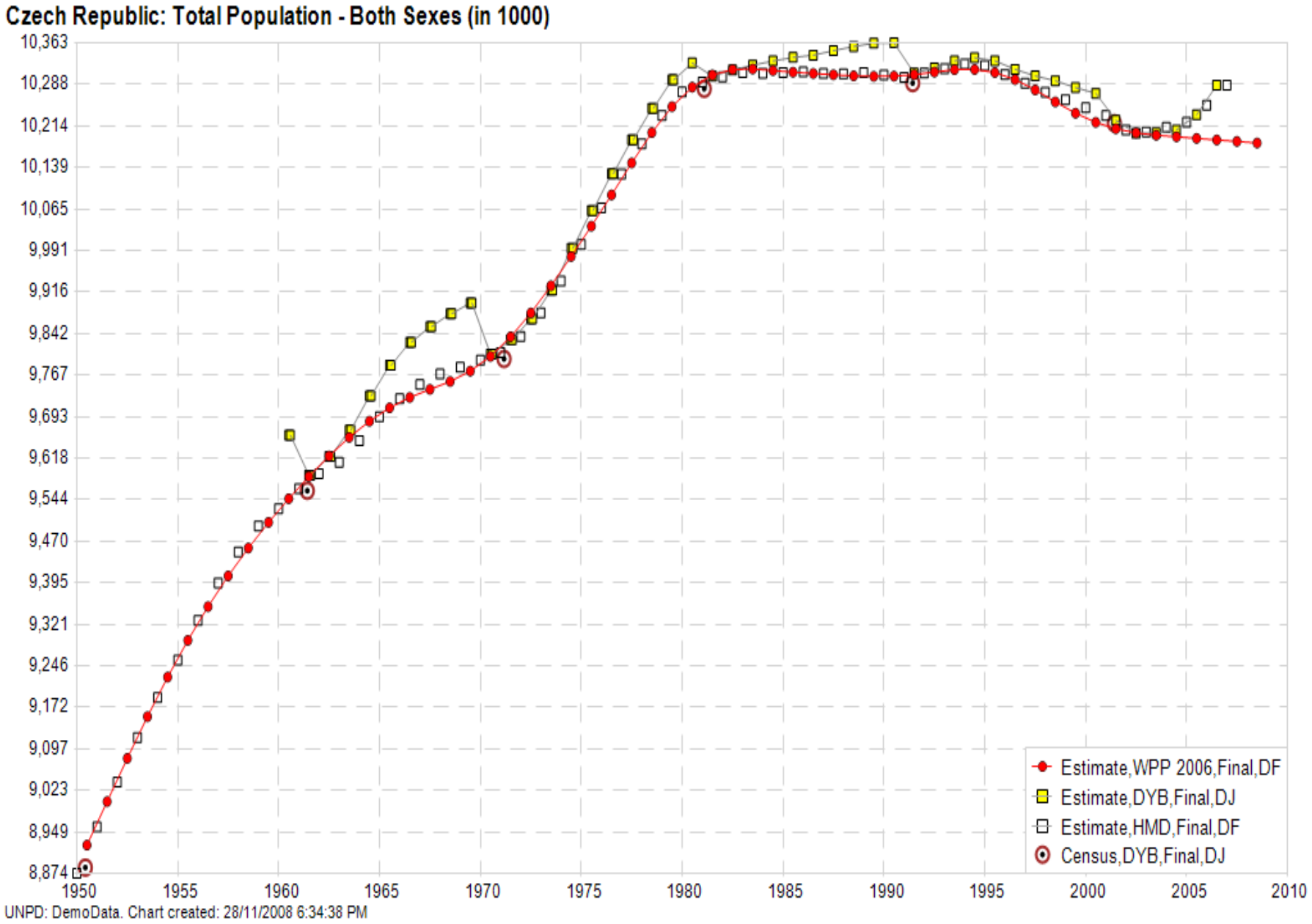


Figure 5: Czech Republic: Retrospective correction of population estimates after census:

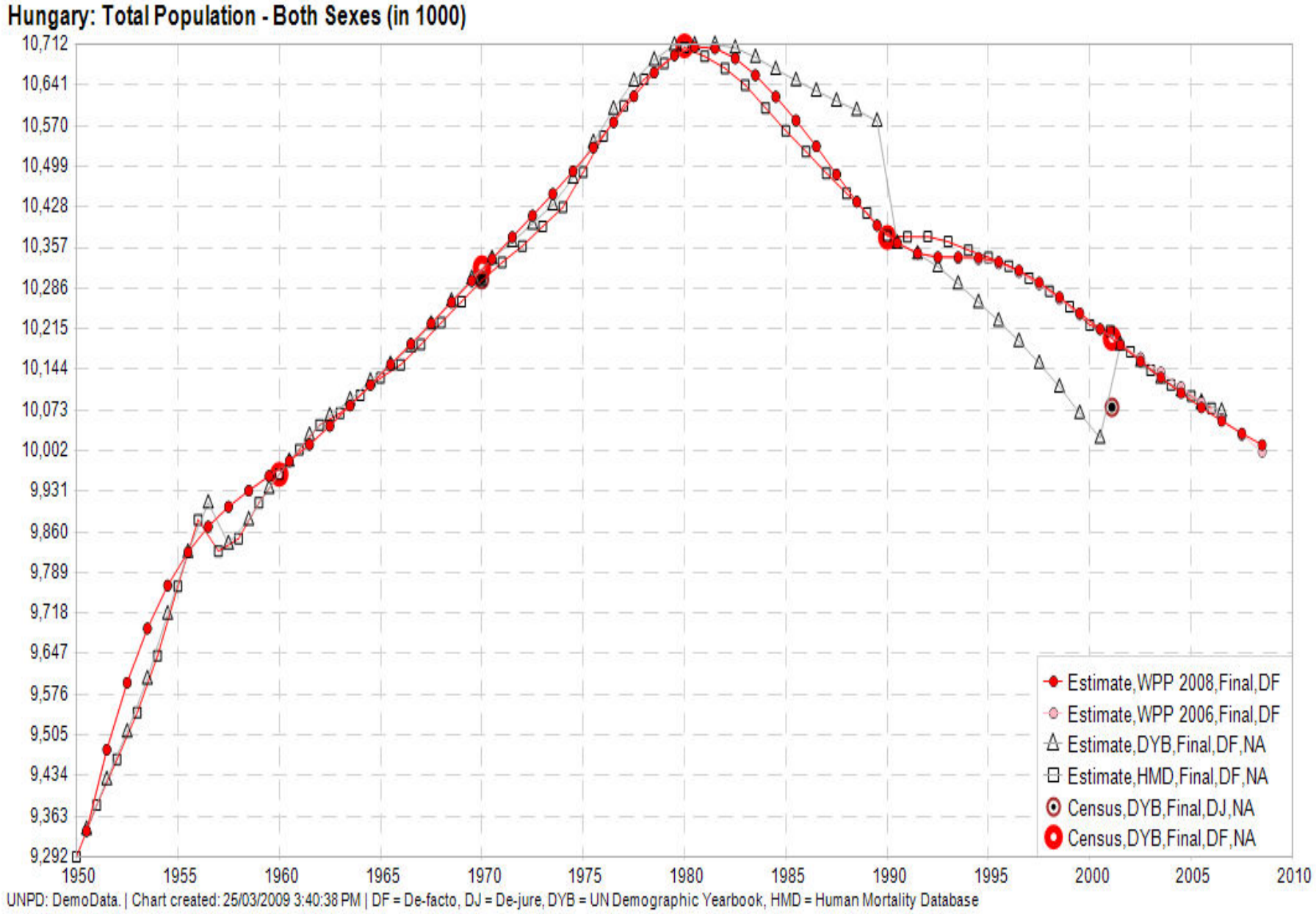


Figure 6: Screenshot of the Interface to the DemoData empirical database: Total Population

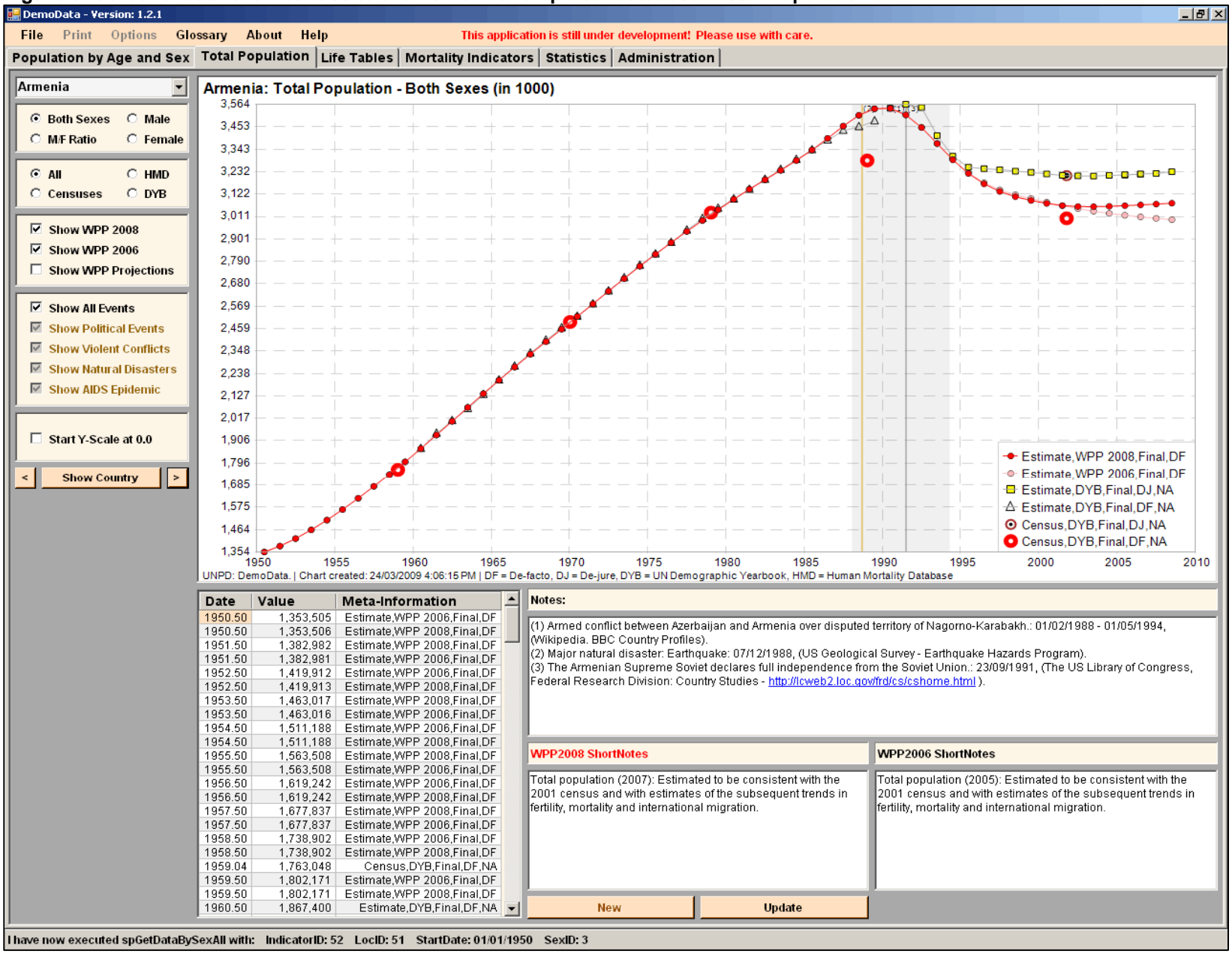


Figure 7: Screenshot of the Interface to the DemoData empirical database: Population by Age and Sex

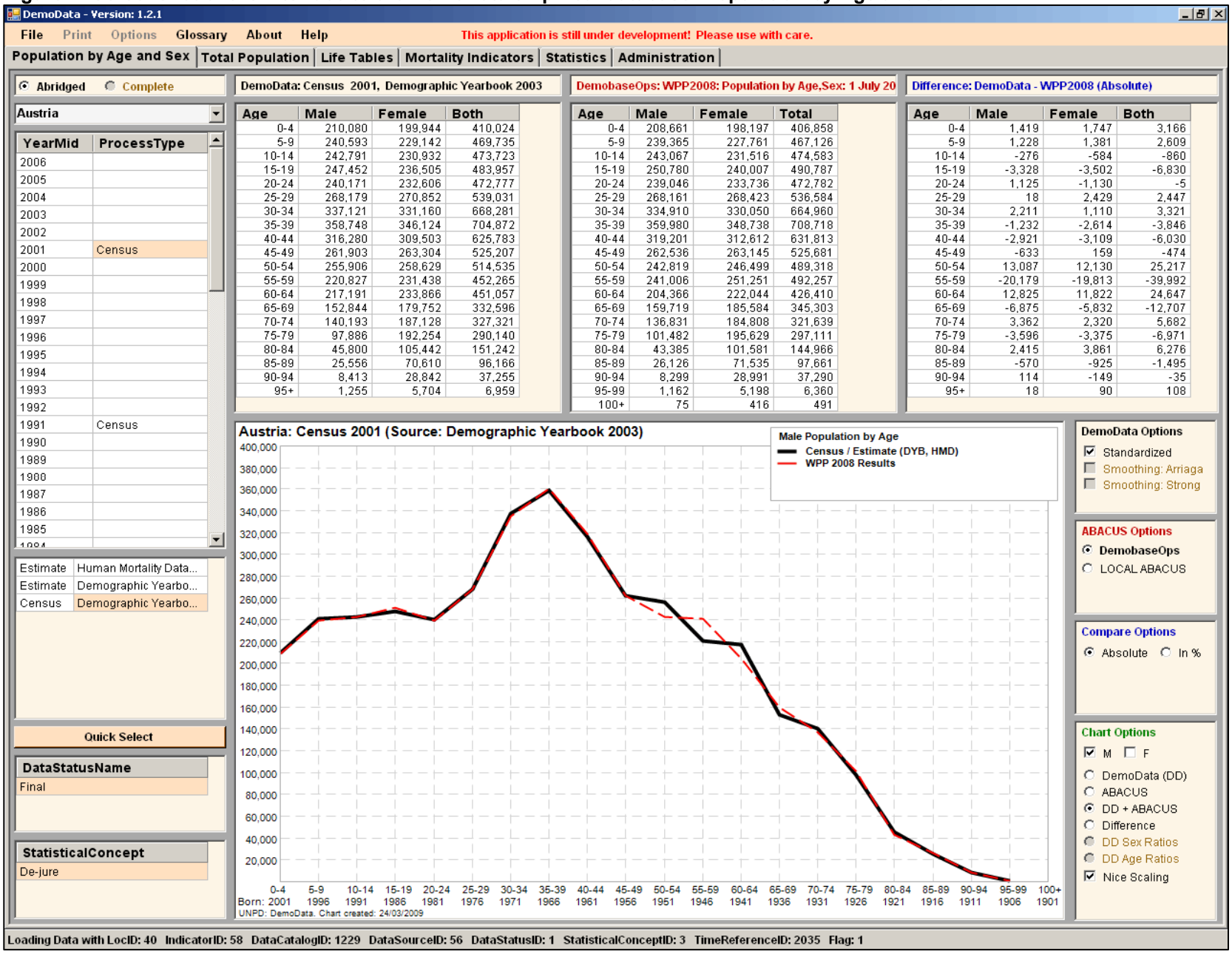


Figure 8: Screenshot of the Interface to the DemoData empirical database: Under-Five Mortality (5q0)