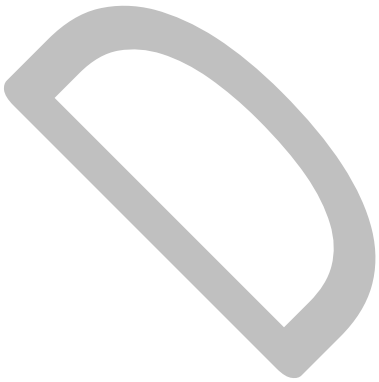
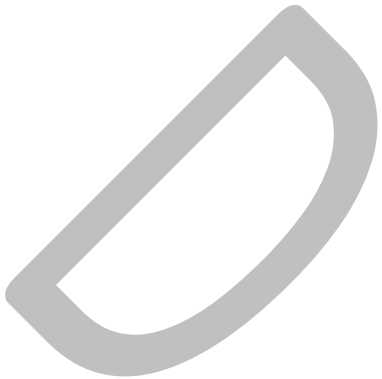




   c. From other sources
4. Census figures by 5-year age groups
   a. From UNSD DYB based on data provided by national statistical authorities
   b. From country national statistical authorities
   c. From other sources
5. Survey data on population by single years of age (usually as percents)
6. Survey data on population by 5-year age groups (usually as percents)

Note that for items 5 and 6, an estimate of the total population, by sex will be needed. January 1 population data from the Human Mortality Database (HMD) project can also be considered or ua12 7910(u)TQW* nBT/F1

[Action item: Compile a more complete set of PES estimates of net coverage errors by age to look at spatial and temporal patterns. This could also include separately identified estimates based on demographic analysis. t UNPD status as of mid-march 2020: an initial data compilation has been completed, and need to be further consolidated.]

PES's have been done in various forms in the US since 1950, but often the sample error is so high that it is hard to determine definitively whether there was an undercount or an overcount.

Age-specific under- or over-enumeration can significantly distort the age distribution of the population. A consistent pattern across many countries is the undercounting of young children. The Infant Enumeration Study 1950 (U.S. Census Bureau, 1953) looked at infants enumerated in the 1950 census to the birth registration data, and followed up on registered births that were not matched by a questionnaire sent to t

Some countries (like Canada) do use the results of the PES to adjust their population estimates. It is important to read the metadata (which may include footnotes in tables) to check on whether estimates are adjusted or not. In the HMD estimates for Canada (2019a):

> Population estimates obtained from census enumerations are subject to undercount errors, which vary from census to census (typically 1 t3% for recent censuses). Starting in 1971, Statistics Canada has been correcting the July 1st population estimates to adjust for the net undercount in the census. No such adjustments were applied to population estimates for the earlier years.

This does raise questions about whether pre-1971 populations should be adjusted to be consistent with later estimates.

PESs taken in India in 1981, 1991, and 2001 show relatively consistent patterns of net omissions by age, with high levels for the population under 5 and for young adults (males 15-39 and females 15-29 in 2001), and some indications of a rise in the oldest age groups (Gerland, 2013). These figures also show a slow rise in overall net omission rates over time from a minimum of 0.68% in 1961 to 2.33% in 2001.

B.2.c. Other sources to assess completeness

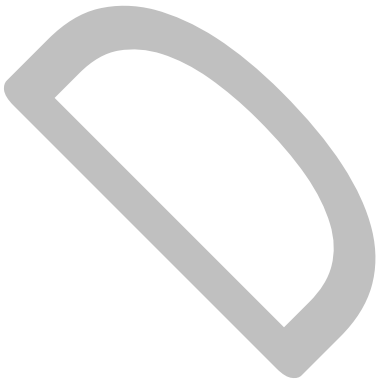
Other sources of data can be used to assess the quality of enumeration. These are mostly administrative sources that are collecting data for other purposes. These can include:

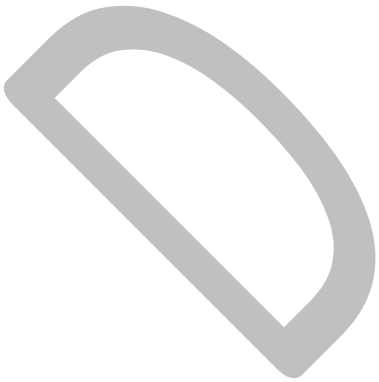
- Vital statistics from birth (and death) registration or from health authorities
- Education statistics (numbers of students by age): particularly relevant for primary and secondary age groups where/when school enrollment is quasi-universal
- Voter rolls for adult voting age groups
- Tax records for adults
- Immunization records, especially for infants under age 1 (e.g., DTP3)
- Records of issuance of IDs to residents
- Old-age pension and medical records

As noted above, registered births and deaths and records of migration flows can be combined to produce demographic estimates of expected populations (either total or by age and sex).

The use of other administrative data tends to be used to provide alternative estimates for particular age ranges of the population. As noted above, these records and data are collected for other purposes, so the first task is to locate these types of data. Some of these data may be found in statistical yearbooks produced by national governments. This may point to the part of the government that collects this data so that additional information and detail may be obtained from the particular government agency website. Other data may be aggregated by international organizations such as WHO, UNICEF, UNESCO, etc.

The second task is to determine what age group they relate to so that comparisons medical6477542.372a72.02p(d)3(et)7





I have used the Beers (ordinary) in the past, and it works well. At the U.S. Census Bureau we used it regularly, often with the Johnson (Stover, et al., 2008) adjustment for age 0, to create the base population used in the Rural-Urban Projection (RUP) program (U.S. Census Bureau, 2014) for projections. The modified version does some smoothing and does not necessarily reproduce the totals. The Sprague method seems to work well also. The monotonic spline has the advantage of allowing varied inputs, including 10-year age groups or 0-4 and then 10-year groups. This can be helpful to force the excess population at ages ending in 0 to be redistributed in both directions (younger and older). It also guarantees there will be no negative populations, which can happen with Beers and Sprague. The DemoTools monoCloseout routine is a blend of Sprague at younger ages with the monotonic spline at older ages to prevent any negative values. We may want to look at the assumptions at the youngest ages: a test run of the monotonic spline using the 1950 WPP 2019 female population for Bangladesh (admittedly not real data for analysis) resulted in much lower estimates for age 0 compared to other methods (at least 8% lower). However this same data did produce negative numbers at older ages for the Sprague and both Beers versions.

These smoothed/graduated populations can then be used for:

1. Cohort comparison
2. Mortality estimation
3. Census evaluation
4. Estimation of 1950 population

[Action item: Comparison of different methods for splitting 5-year ages into single years of age to determine which works best in certain circumstances.]

[Recommended steps: Single age data should be used as much as possible and be smoothed if necessary. The original or smoothed single age data should be compared to graduated 5-year data. Series should be compared, and a set of conditions/criteria/threshold be formulated for when to accept one series over another.]

B.5.b. Ageoverstatement or understatement

Age over- or under-statement (other than related to age heaping) can be more difficult to correct. The LAMBdA (2019) model discussed above dealt with this for older ages (above age 45) by using a model based on conditional net overstatement probabilities and age-specific propensities to over-state age estimated based on the results of a 2002 evaluation study linking Costa Rica 2000 census records to voter register records that included age information from birth certificates. The net overall tendency to overstate

older ages than typically available from censuses (Wilmoth, et al., 2019, Condran, et al.,1991, Jdanov, 2015).

B.6. Comparing population estimates by age and sex

Once the initial analysis of the population estimates is completed, the results based on different censuses (or data sources) should be compared to makes sure that they are consistent. The PAS workbook GRPOPYB.xls (or NewPAS versions GRPOPYBNEW.xls and POP1COHNew.xls) plot populations estimates for different dates by birth cohort instead of by age. This type of plot can also be applied to the reported populations to help identify cohorts that may be smaller or larger than expected.

Figure B-3. Male Population of Djibouti by Census and 5-Year Cohort

If some of the estimates have been based on smoothing (and adjusting) populations by single years of age and some have been based on smoothing and splitting data by 5-year age groups, then this can help identify if there are inconsistencies in the resulting estimates. This type of plot can also help identify cohorts that may have been affected by migration.

Another way of assessing the consistency of the estimates is to compute census survival ratios. Using populations by single years of age should make it easier to do this analysis if the intercensal interval is not 5 or 10 years. Survival ratios are expected to decrease monotonically with age, and to always be lower than 1. In addition, any abnormal deviations in some age groups are expected to follow the birth cohorts as they age if these deviations reflect real historical events

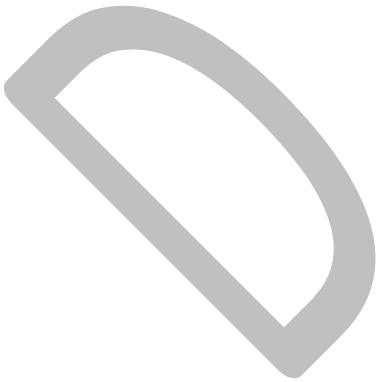
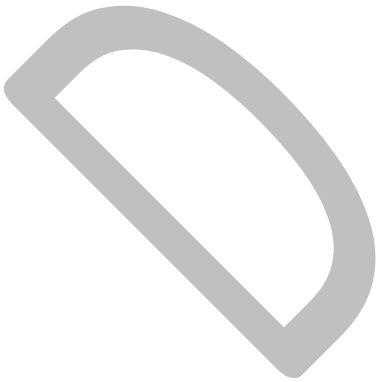
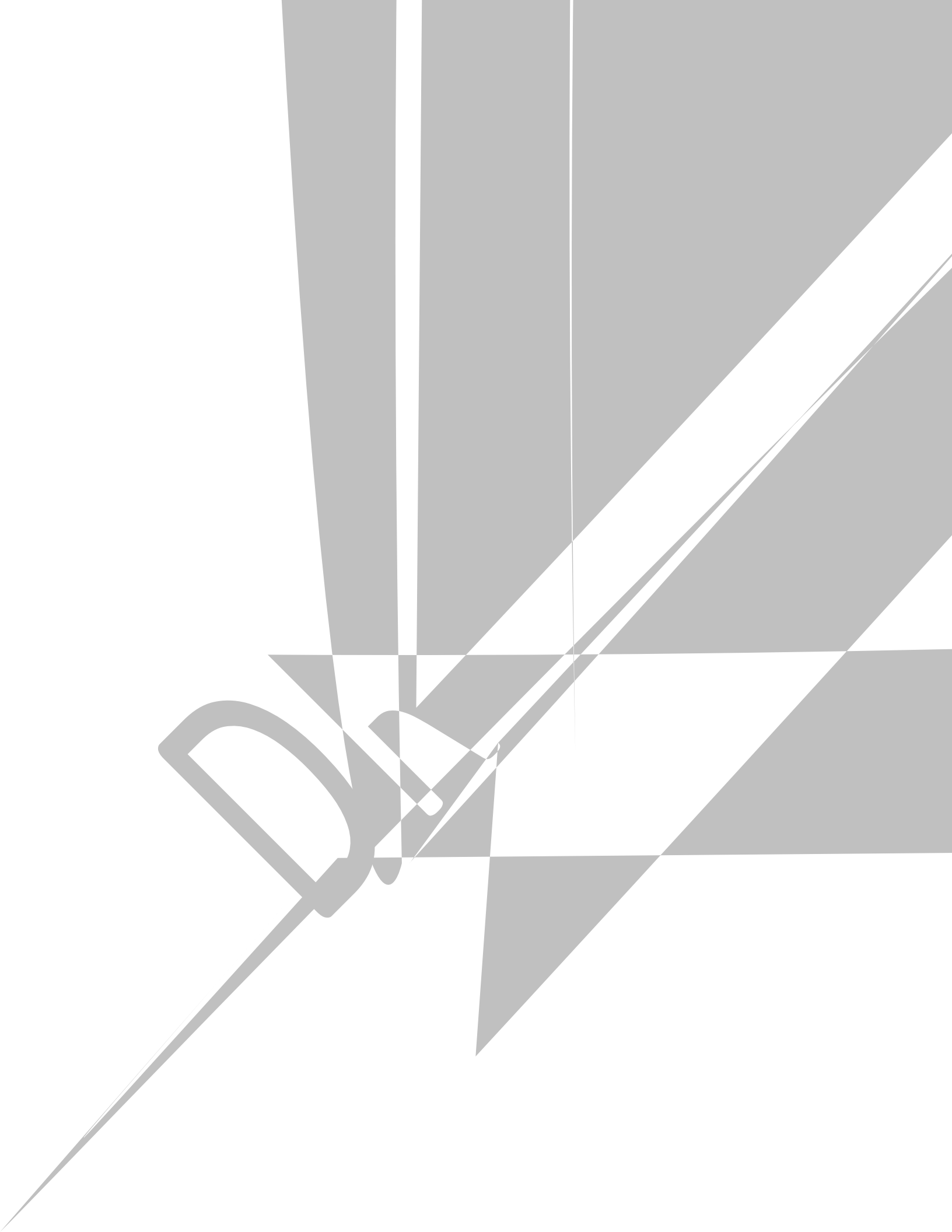
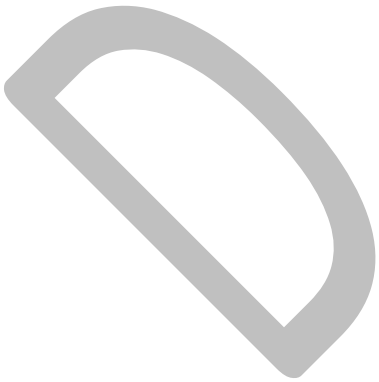
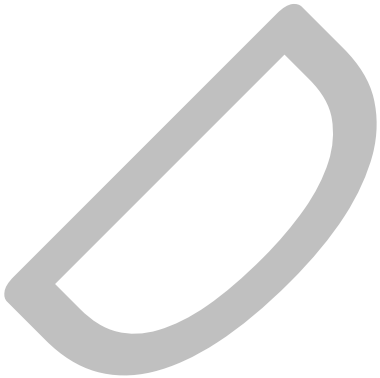


Figure D-1 shows TFR estimates from DHS surveys in D5stimates 0 G88897.94 3F1 11.04 Tf1 0 0 1 300.6 52.104 Tm0 g0 G









H. REFERENCES

Abel, G.J. (2019). Update for WPP2019. https://guyabel.com/publication/bilateral-international-migration-flow-estimates/

Abel, G.J., and Cohen, J.E. (2019). Bilateral international migration flow estimates for 200 countries. Scientific Data, 6:82, 575-594. https://www.nature.com/articles/s41597-019-0089-3.

Azose, J.J. and Raftery, A.E. (2019). Estimation of emigration, return migration, and transit migration between all pairs of countries. Proceedings of the National Academy of Sciences of the United States of America (PNAS), 116:1, 116-122ETQ E12 0q0.9Qq0.00000912 0 612 792 re.04 Tf1 0 0 1 275.64 562.97 T

Johnson, P.D., and Way, P.O., (2003). The Role of Population Estimates and Projections in the Evaluation of Censuses. International Programs Center, Population Division, U.S. Census Bureau. Paper

Preston, S., and Bennett, N. G. (1983). A census-based method for estimating adult mortality. Population Studies 37 (1), 91-104.

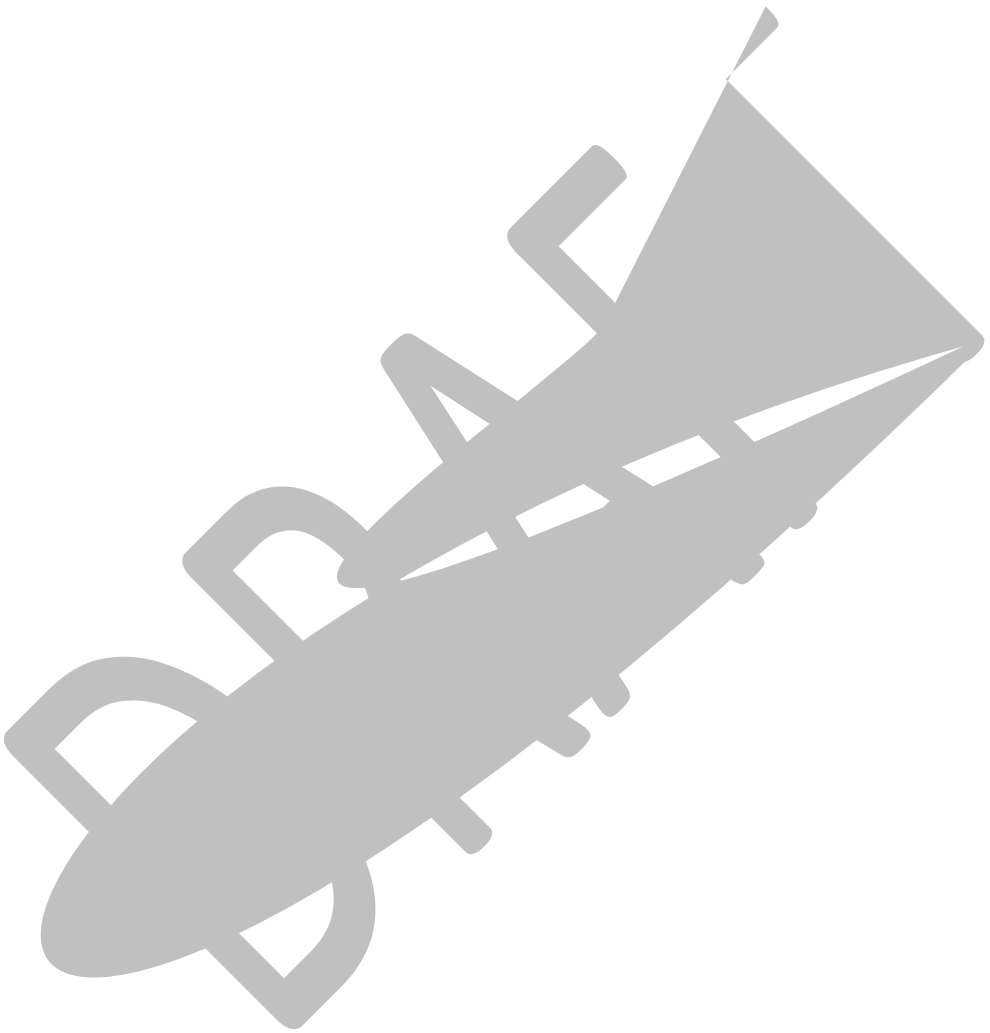
Preston, S., Coale, A. J., Trussell, J. and Weinstein, M. (1980). Estimating the completeness of reporting of adult deaths in populations that are approximately stable. Population Index 46 (2), 179-202.

Preston, S., and Hill, K. (1980). Estimating the completeness of death registration. Population Studies 34, 349-366.

Pullum, T. W., Croft, T., and MacQuarrie, K. L. D. (2018). Methods to Estimate Under-15 Fertility Using Demographic and Health Surveys Data. DHS Methodological Reports No. 23. Rockville, Maryland, 144.20i24(A)-7(:)7( )9(I)-8(C)-9(Fi24(.)] TJETQq0.00000912 0 612 792 reW* nBT/F1 11.04 Tf1 0 0 1 196

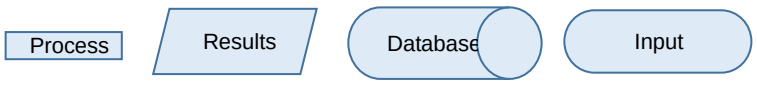
United Nations. Department of Economic and Social Affairs, Population Branch. (1956). Manual III. Methods for Population Projections by Sex and Age. Population Studies No. 25. ST/SOA/Series

United Nation:NÍ( )] TJETQq0.00000912 0 612 792 reW* nBT/F1 11.04 Tf1 0 091 329.05 6335.26(d)3 108.067ewyY0



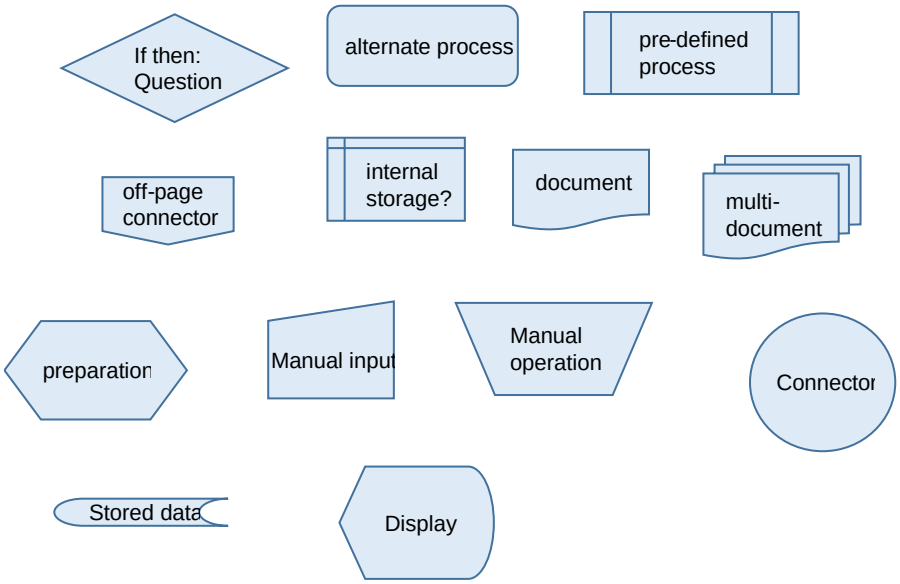
Legend for symbols used:

GBD symbols



Process flow

Other symbols to consider



Data flow

Grouped items

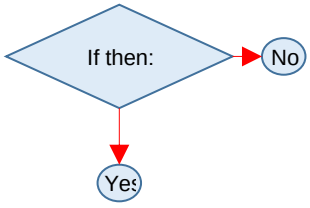
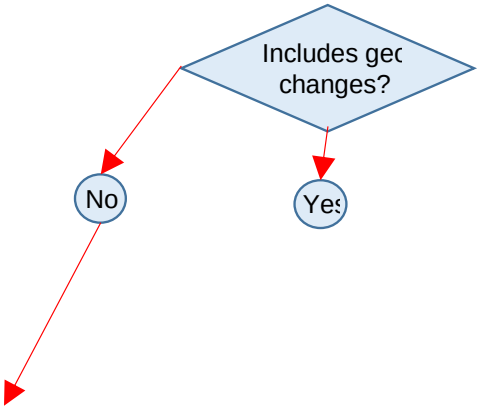
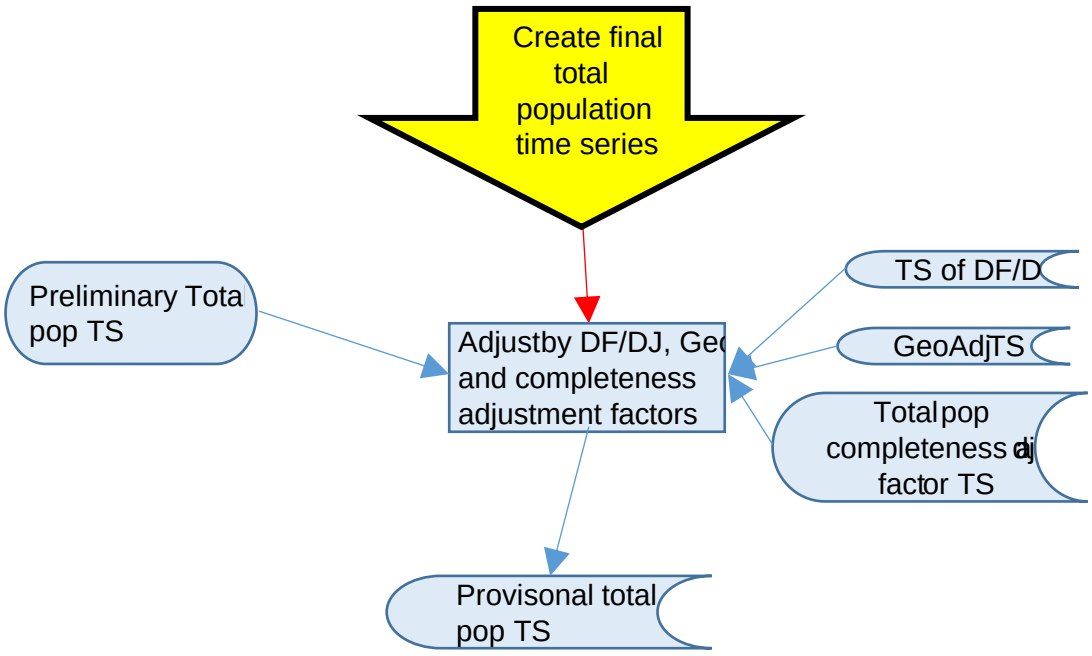


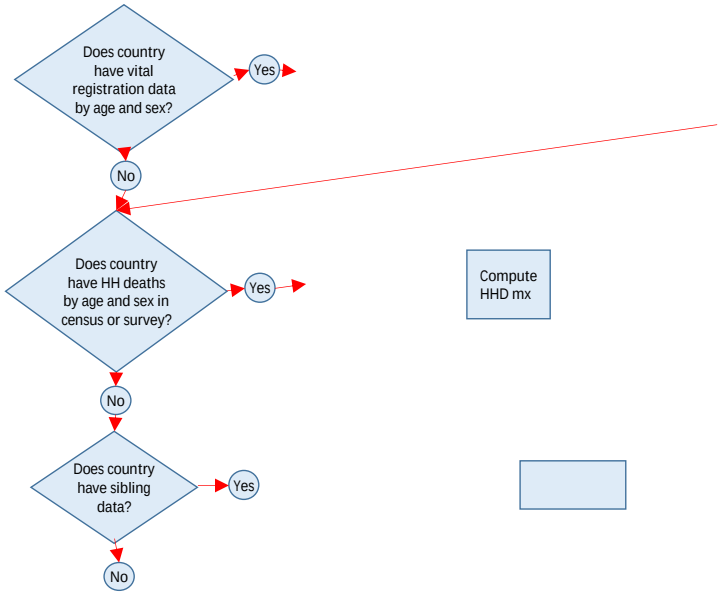
Diagram PopT-2: Time Series of Geographic Adjustment Factors

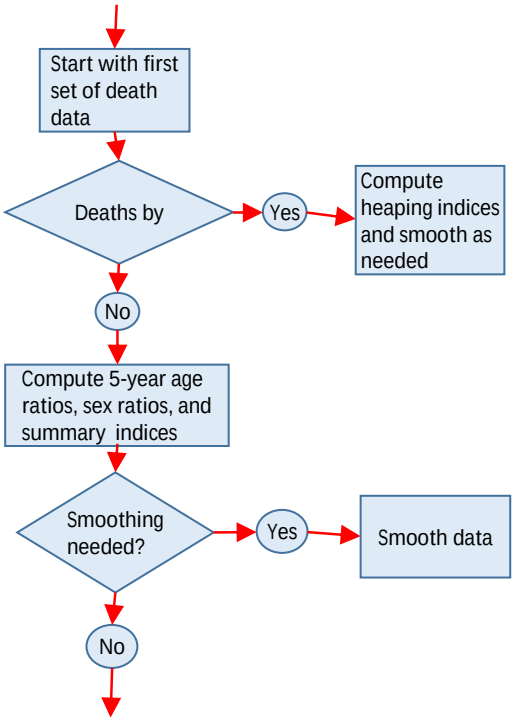


Create

Diagram PopT-4: Final Time Series Adjustme




Does country have vital registration data by age and sex?
Yes
No
Does country have HH deaths by age and sex in census or survey?
Yes
Compute HHD mx
No
Does country have sibling data?
Yes
No



1. What to do with high sex ratio indices of deaths by age?
    a. May indicate differential mortality patterns by sex (like changes in Colombia)
    b. May indicate differential levels of completeness of death registration
    c. Maybe don't bother looking at sex ratios (except to make sure there aren't any sudden changes that may point to data errors or special circumstances). Sex ratios of mx values are usually more informative.
2. Check on cohort patterns?
    a. May indicate abnormal birth cohort size (baby boom or bust)
    b. Possible long-term impacts of events in early childhood?

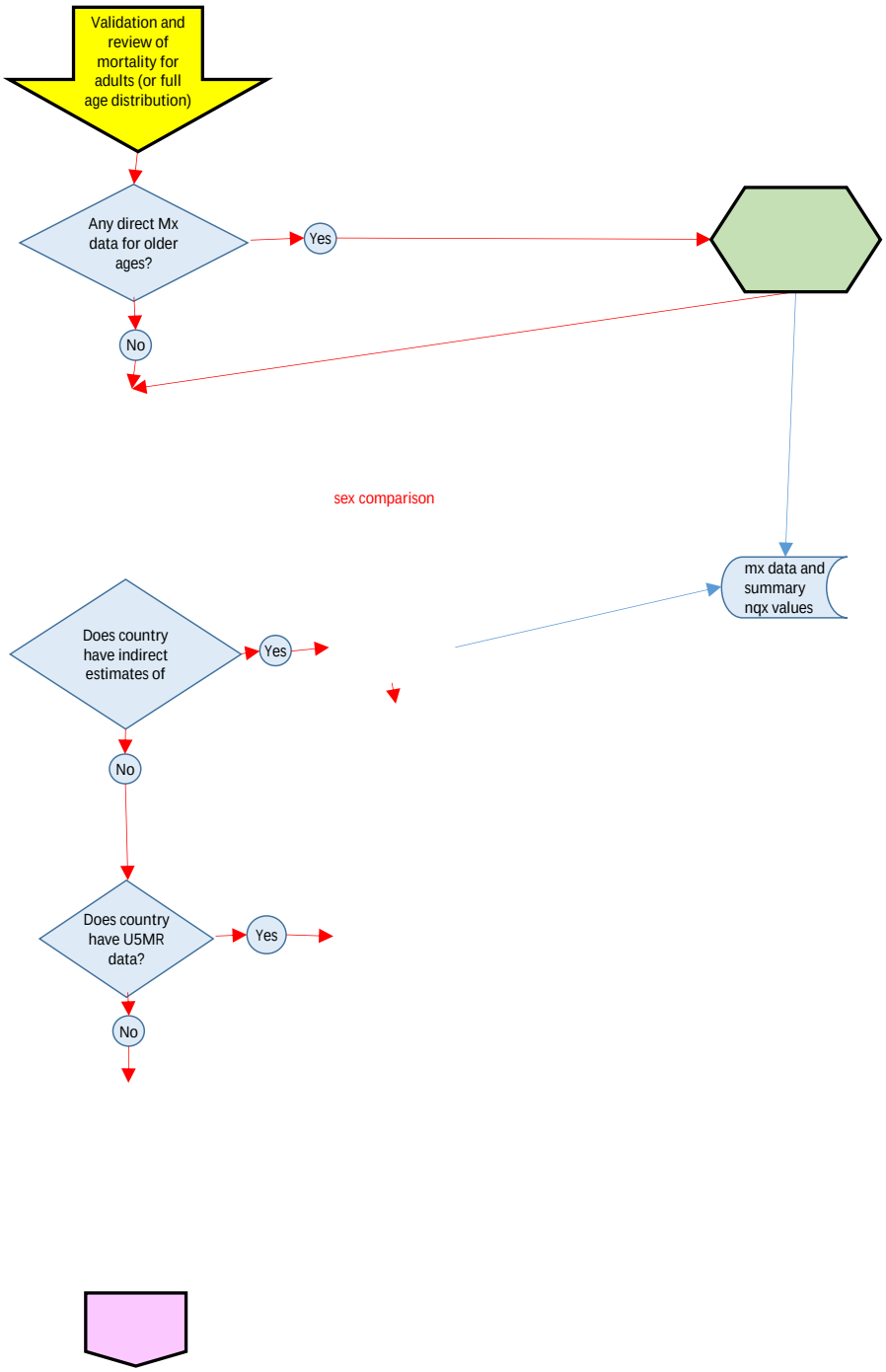

Validation and review of mortality for adults (or full age distribution)
Any direct Mx data for older ages?
Yes
No
sex comparison
mx data and summary nqx values
Does country have indirect estimates of
Yes
No
Does country have U5MR data?
Yes
No

Diagram MortC-1: Child Mortality Overview

Diagram MortC-3: IGME Adjustments and B3 Model

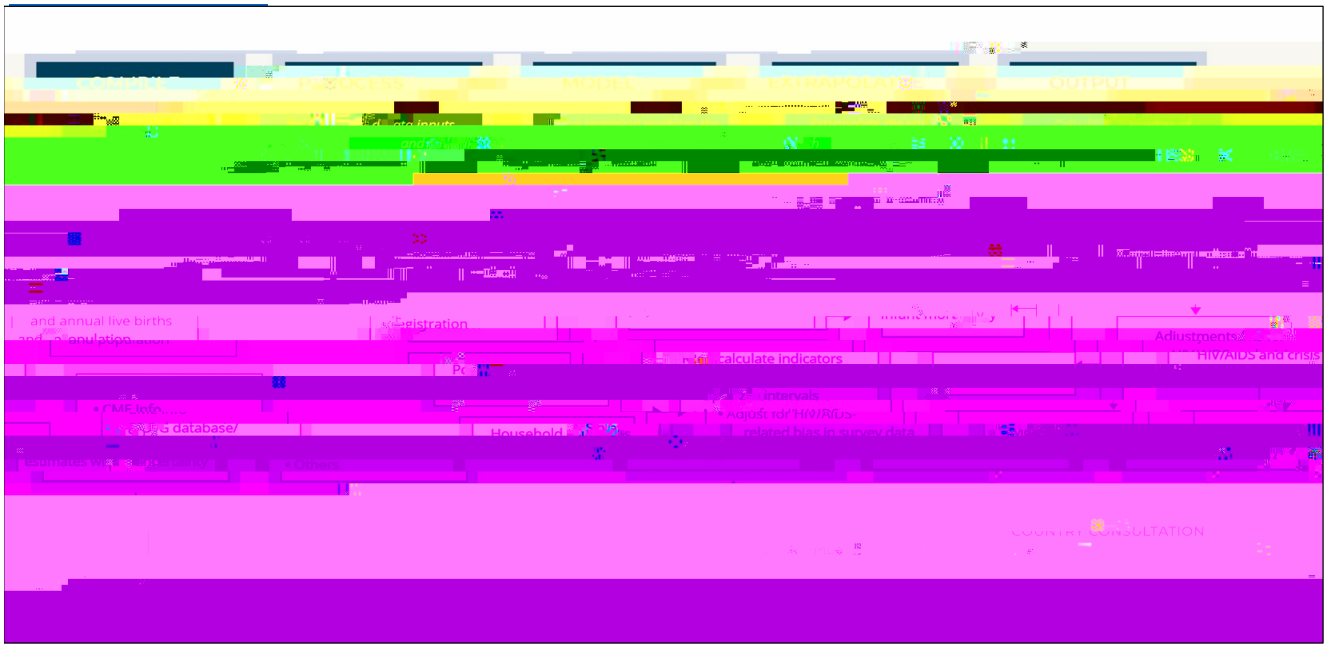


Diagram MortC-4: IGME Notes