

IPUMS-Europe Census Microdata Harmonization Project

Overview (revised: September 1, 2004) <http://www.hist.umn.edu/~rmccaa/ipums-europe>
dissemination site: www.ipums.org/international (harmonized data for France & other countries)

Welcome! Thanks to the support of official statistical agencies in 18 European countries (see list of partners below) and major funding by the National Institutes of Health, one of the world's largest, integrated scientific instruments for the study of human populations is now under construction. The database will contain anonymized microdata samples, encompassing as many as 50 censuses and totalling more than 70 million person records. Researchers already are using, free of cost, integrated microdata for 7 countries: China, Colombia, France, Kenya, Mexico, USA and Vietnam from the IPUMS web-site. Brazilian microdata (5 censuses) are undergoing final testing and will be placed on the main site in the near future (until then, please see <http://beta.ipums.org/international> for a preview).

Constructing the European database will be easier than what many imagine. From the IPUMS-International work completed so far, we have learned that many country projects can be underway simultaneously without the need for a great deal of travel or consultation between them. Many problems are readily resolved by email. The Latin America project, which began last year and serves as a prototype for IPUMS-Europe, may be viewed at: www.hist.umn.edu/~rmccaa/ipumsla

Please note that computer work will be performed at the Minnesota Population Center (MPC) using software tools already developed from working on the first eight countries. The project has 4 phases: collection, preparation, harmonization and dissemination. With each National Statistical Institute (NSI) partner, advancing from one stage to the other is determined by when the work is completed:

1. **Collection Phase** ("year 1"): See below: "Billing Instructions", "Documentation Needed" & "Microdata Specifications". Please send these materials via courier service.
 - NSI: Provide copies of original documentation and microdata for each census
 - MPC: Pay license fees to NSI; recover endangered microdata
 - Centre d'Estudis Demogràfics (CED, Barcelona): Workshop, July 25-28, 2005
2. **Preparation Phase** ("year 2"): Country-specific details will be provided once we have examined the country's census documentation and microdata.
 - MPC & CED: Scan and translate documentation as necessary; validate and clean microdata, confirming record structure and contents (work to begin as soon as received).
 - NSI or in-country expert consultants: Provide methodological assistance with anonymization issues and difficult country-specific variables, such as administrative geography, economic and educational variables, etc.
3. **Harmonization Phase** ("year 3"):
 - MPC & CED: Design integration tables for each variable (original code, harmonized code); write metadata for users explaining comparability strengths and weaknesses
 - NSI, or in-country consultants: Test alpha site for methodological coherence and consistency
4. **Dissemination Phase** ("year 4" and beyond):
 - MPC & CED: Provide harmonized extracts to researchers; conduct workshops at regional or international meetings to promote use by national researchers; provide NSI partners with usage data and publications resulting from research using integrated microdata