

Composition of Religious and Ethnic Groups Project (CREG) v.1.02

Variable Names for CREG Ethnic and Religious Groups Datasets

Long Datasets

Variable	Description
Country	Country
Cowcode	Correlates of War (COW) Code ¹
Year	Year
Independent Country?	Gleditsch and Ward Independence Dummy Variable
Group Name	Name of group
Group Proportion (Raw Data Point)	Raw population data (proportion of population)
Group Estimate	Population estimate (proportion of population)
Outlier	Raw population data marked as an outlier
Anchor	Estimate used to "anchor" a population estimate

Wide Datasets

Variable	Description
Country	Country
Cowcode	Correlates of War (COW) Code ¹
Year	Year
Independent Country?	Gleditsch and Ward Independence Dummy Variable

Note: Remaining group name variables contain each group's proportion of the population in a given year.

¹ Changes were made to the COW Codes used in this analysis. Please read the "Notes" section to learn more about the changes made.

Notes

General Notes

1. **Missing Countries²:** There were insufficient data to make ethnic group estimates for **Cameroon, France, India, Kosovo, Montenegro, Mozambique, and Papua New Guinea.** There were insufficient data to make religious group estimates for **Democratic People's Republic of Korea, People's Republic of China, and Republic of Vietnam.**
2. **Country identifiers (COW codes):** The values used were based on those in Gleditsch and Ward's list of independent states with two exceptions, post-USSR Russia and Post-Yugoslavia Serbia. The changes in population data between the pre-breakup period and the post-breakup period were large enough that a separate country had to be created to make post-breakup estimations. The following COW codes were used:

Country	Cowcode	Years
Serbia	340	1992-2013
Russia	393	1991-2013
USSR	365	1945-1991
Yugoslavia	345	1945-1992

The Gleditsch and Ward coding scheme for post-unification Germany and post-unification Yemen were used:

Country	Cowcode	Years
Federal Republic of Germany	265	1945-2013
Yemen (Arab Republic of Yemen)	678	1945-2013

3. **Ethnic Group Labels:** Ethnic group labels are consistent across all countries with the following exceptions:
 - Spaniard – used solely in Spain. All references outside of Spain use the label “Spanish.”
 - Han Chinese – used solely in China. All references outside of China use the label “Chinese.”

² The countries included in this data are from the Cline Center's Societal Infrastructures and Development (SID) Project which include all countries that were independent in 2004 and whose populations were greater than 500,000. To learn more about the SID Project please visit: <http://www.clinecenter.illinois.edu/research/sid-project.html>.

Changes Made Over Time with CREG

v.1.02 (5/12/2014)

1. Groups labeled as “Hungarian” were changed to “Magyar.”
2. Ethnic group estimates for France were removed due to modelling issues. Hope to have ethnic group estimates for France in later editions.

v.1.01 (5/01/2014)

1. Added ethnic group estimates for the following countries:
 - Argentina
 - Belgium
 - Poland
 - Republic of Korea
 - Rumania
 - Sudan.
2. In addition, United Arab Emirates had ethnic group estimates for 2014 and 2015. These estimates were removed for consistency.