

Codebook for Documenting Global Ethnic Disparities (3/20/2020)

Codebook. The codebook is organized into several sections.

- Language. These codes will give information about the status of the ethnic group's primary language and the privilege experienced by an ethnic group by having the state language as their primary language.
 - STATELANG. Is the language primarily used by members of ethnic group today a state language in this country? This only counts for languages that are designated for use in official state affairs and used in official state documents. In the comment box, please indicate the State language(s) as well as the primary language of the ethnic group. What counts as a distinct language (e.g., Serbian vs. Croatian) will depend on what are the most salient distinctions in the specific country. If the ethnic group has shifted to primarily using the state language in recorded history, code Y and leave a comment. Y/N
 - RECLANG. Is the primary language of the ethnic group an officially recognized language or dialect in this country? This should be only coded as N, in those cases where there is specific evidence that the group's language is not recognized. Y/N
- Ethnicity characteristics. These variables are intended to give some idea of where the ethnic group name came from. It is not directly about privilege or discrimination, but provides useful information about society's define ethnic groups.
 - LANGUAGE_DEMONYM. The ethnic designation is synonymous with a specific language or dialect spoken commonly by members of the ethnic group (e.g. French, Aymara)? It is not necessary that members of the ethnic group in the specific country under consideration currently speak the language (e.g. Germans in Bolivia). It's also OK if other ethnic groups speak the language. Exclusions: *ethnic names like "Indian" or "American" are not synonymous with specific languages and so shouldn't be counted.* Y/N
 - COUNTRY_DEMONYM. The name for the ethnic group is a **demonym (see definition below)** that is synonymous with a nation-state (e.g. French, Rwandan, Yugoslavian, Indian)? This should be a nation-state that was recognized at least once by the U.N. in the last 50 years. It is not necessary that the ethnic group is currently living in the country (e.g. Indians living in South Africa would count here). Exclusions: *This should not include Arabic, which is defined by a common language and culture, but not necessarily by coming from Saudi Arabia. This should not include cases where the ethnicity name specifically refers to speakers of a language (e.g. "English-speaking", "French-speaking"). This should not be used when the country name appears in a composite (e.g. "Afroboliviano").* Y/N
 - AUTONOMOUSZONE_DEMONYM. In the last 100 years, there has been a geographically bounded autonomous zone within or across other countries that bears the name of this ethnic group and is in some sense controlled by members of the ethnic group (e.g. Tibetan, Kurdish). Exclusions: *This should not be coded if it is already coded for a country demonym.* Y/N
 - REGION_DEMONYM. The name for the ethnic group is a demonym based on region of the world larger than a nation-state (e.g. European, Afrodescent, South Asian)? Exclusions: *This should not include Arabic, which is defined by a common language and culture, but not necessarily by coming from the Arabian Peninsula. This should not be coded if it is already coded for a country or autonomous zone demonym.* Y/N

- PHYSICALFEATURE. The name of the ethnic group is based on some physical feature common among members of this ethnic group (e.g. Black, White, Colored). *Exclusions:* This should not include names that imply a regional origin (e.g. African-American, Euro-American), unless they also make specific reference to physical features.
- PASTNATION. This ethnicity was the dominant ethnic group or descends from the dominant ethnic group in a nation/empire/kingdom with a centralized government (e.g. Songhai, Malinke)? The past nation does not need to have been in the current country. The ethnic group's name should be derived from the nation's name. This should only be applied to situations where any kingdom, nation, or state associated with this ethnic group is no longer in existence. Y/N
- MIXTURE. The name of the ethnicity explicitly implies a mixture of two groups (e.g. Mestizo, Creole, Mulatto)? If this is coded, also code for REGION when relevant. Y/N
- OCCUPATION. The ethnic group's name is defined by an occupation. This is a common way of defining castes in South Asia. It is not necessary for members of this ethnic group to currently be engaged in this occupation. Y/N
- Ethnic recognition. In only rare cases are entire ethnic groups deprived of citizenship in a country, but we record this for those rare cases.
 - CITIZEN. Are members of the ethnic group officially recognized as citizens or legitimate residents of the country? (e.g. Myanmar does not recognize Rohingya as an official nationality, Roma are not officially recognized in many eastern European countries). This should be only coded as N in those cases where there is no specific evidence that the group is not recognized. Simply mention of "secondary citizen" does not count here, because they are still citizens unless additional evidence is provided. Y/N
- Religion. This provides information about whether the religion of the ethnic group is the privileged religion in a country.
 - RELIGION. Is the currently dominant religion of the ethnic group in this country also the dominant religion in the country? The dominant religion means the religion practiced by the most people. What counts as a distinct religion (e.g., Protestant vs. Catholic) will depend on what are the most salient distinctions in the specific country. This should count as Y even if the religion is a mixed religion (see below). In the comment box, please indicate the country's dominant religion as well as the dominant religion of the ethnic group. If the ethnic group has shifted to primarily this religion in recorded history, code Y and leave a comment.Y/N
 - RELIG_MIX. Is the currently dominant religion of the ethnic group in this country a mixture of traditional beliefs/practices and the dominant religion in the country? This means the religion is a mixture of practices and influences and not simply that some people practice one religion and others practice another. Both RELIGION and RELIG_MIX can be Y if applicable. Y/N
- Secession
 - SECEDE. Has a substantial section of the ethnic group mounted an organized attempt to secede from the current country to form an independent state? This does not need to be successful. Y/N
- Self-governance

- SELFGOV. Have members of the ethnic group established at least partial independent control over land/politics/community services in their territory that is officially and explicitly recognized by the national government? This needs to be current control. If this is just about education in the local language code as N and make a comment. If members of the ethnic group have simply tried to gain such independence, this code does not count (instead consider SECEDE). Y/N
- Political power
 - HEADSTATE. Has a member of the ethnic group been a head of state for that country in the last 30 years? This can include heads of state that are democratically elected, or who come to power by other means (e.g. military coup). Examples are president, monarch, prime minister. In some cases there may be more than one individual that fits this definition at one point in time (e.g. queen and prime minister). Include if the leader has recognized ancestry from this group. Only count individuals, not ethnically defined political parties. This should only include heads of state for the entire country, but other kinds of leaders can be added to comments. *Please include name and year of leader in the comment box.* Y/N
 - COLONIZER. Did this ethnic group forcefully colonize the region in the recorded past (even if it was before the founding of the current country)? This also counts if the ethnic group gained a powerful position in society because it was an ally of the colonizer. See definition of **colonization** below. If the ethnic group is a mixture of past colonizers and colonized populations, then put Y for both COLONIZED and COLONIZER. Y/N
 - COLONIZERTIME. If so, when did this occur? Give the date range
 - COLONIZED. Was this ethnic group forcefully colonized within the region of this country in the recorded past? See definition of **colonization** below. Y/N
- Slavery
 - SLAVERY. Were members of this ethnic group systematically enslaved within the region of this country by members of OTHER ethnic groups? This can include forced labor (but please add a comment for this). If members of this ethnic group were systematically enslaved by members of a specific other ethnic group, please record this in the comment section. This can include forced labor at any time in the recorded past. It can also include forced labor due to outsiders colonizing the specific country and also members of the ethnic group targeted for capture to be sold as slaves in other countries. If it is not clear whether this happened within the borders of the country, then code as Y and make a comment. This should not include cases where members of the ethnic group were enslaved by others within the same ethnic group (see SLAVECASTE for this). Y/N
 - SLAVERYTIMEFRAME. If so, when did this occur? Give date range
 - ENSLAVER. Did members of this ethnic group (or cultural ancestors of this ethnic group) systematically enslave others within this country? If they enslaved members of a specific other ethnic group, please record this in the comment for this section. Y/N
 - SLAVETRADER. Were members of this ethnic group (or cultural ancestors of this ethnic group) described as systematically engaged in the capture and sale of slaves? If they enslaved members of a specific other ethnic group, please record this in the comment for this section (and also the time period that this occurred during). Y/N
 - SLAVECASTE. Did this ethnic group include a caste system in which one of the castes is defined as a “slave caste”? Y/N.

- Migration
 - FORCEDAWAY. Did substantial numbers of the ethnic group move either within or away from the country to avoid ethnic-based persecution or discrimination in the recorded past? This can include forced evictions, religious persecution, or ethnic cleansing, as well as concerted attempts to make the ethnic group move even if this is not due directly to the ethnic identity of the group Exclusions: *Do not code if the only reason for moving is because the group is nomadic or because the group is moving for economic reasons unrelated to their ethnic membership.* Y/N
 - FORCEDAWAYTIME. If so, when did this occur? Give the date range
- Genocide
 - LETHALVIOLENCE. Has the ethnic group been subjected to **genocide** (see definition below) or targeted lethal violence within the country in the last 100 years? This can include: (1) state-sponsored killings that are ostensibly for political reasons but that differentially affect some ethnic groups over others, (2) political violence where deaths are concentrated within an ethnic group, (3) mass killings that uniquely affect a large number of members of a specific ethnic group, (4) “tribal conflict” where one ethnic group is subjected to lethal violence from another, (5) targeted killings based on land disputes and in response to rebellions. This should not include deaths due to infectious disease unless there is evidence that one group purposely caused this (e.g., smallpox blankets). Please record in the comments, the approximate number of individuals killed. Y/N
 - ETHNOCIDE: Has the ethnic group been subjected to **ethnocide** (see definition below) within the county in the last 100 years? This can include, punishing members of an ethnic group for engaging in cultural, religious, or linguistic practices characteristic of the ethnic group. This can include examples of an ethnic group having their traditional means of livelihood taken away from land grabs or destruction of their environment. This can include government efforts to change cultural practices even if these are objectionable to the coder. EXCLUSIONS: Simply being forced from land or experiencing repression without clear evidence that it systematically destroyed the culture or language of the group should not be counted here. Language loss that is not a result of a concerted effort should be excluded. Exclude simply mention of laws restricting immigration of members of this group. Y/N
- Discrimination
 - DISCRIMINATION. Have members of the ethnic group consistently been discriminated against because of their ethnicity in the boundaries of this country. This can include
 1. limitations on what jobs they can take, discrimination in the job market, different pay of the same job, restrictions on what kinds of land and material they can own, selective jailing, and exclusion from public services, such as schools, hospitals, bathrooms, drinking fountains, beaches, or from mixing with others in the country (e.g. laws against inter-ethnic marriages).
 2. laws restricting immigration of this group from entering the country.
 3. members of the ethnic group voicing concern about discrimination against the ethnic group.
 4. Official documents using names for the ethnic group that are not preferred or wanted by members of the ethnic group.
 5. This can also include cases of ethnocide

6. Please also include other unlisted forms of discrimination of you feel they are relevant.

In the comments, please make sure to list the specific details of discrimination.

EXCLUSIONS: If discrimination occurs *within* the ethnic group, mark N, but make a note in the comments about the evidence. Do not code this if there is only evidence of greater poverty, lower schooling or lower employment among this group without explicit information about ethnicity-based discrimination. Do not code this if it is simply about being forced to move to new locations. Do not code if this occurred to the ethnic group in a neighboring country (but please comment if necessary). Do not code if this occurred only during temporary occupation by an outside country. Y/N

- Dataquality. This code is a general assessment of quality of data available for this ethnic group relevant to the codes above. Unlike the codes above, it's a scale from 0 to 3 with the following meanings. Please simply use the numbers 0,1,2,3 when coding this.
 - 0: No information found
 - 1: Limited information found from personal/informal sources (e.g. blogs)
 - 2: Limited information found from solid credible sources & peer reviewed journals
 - 3: Ample information found from credible sources & peer reviewed journals

Definitions

Colonization: A foreign ethnic group establishing a government in a new region that exerts social control over the existing population and resources of a region. Although colonization might involve migration of members of the foreign ethnic group to the new region, migration is not sufficient to count as colonization. Territorial invasions during wars (e.g., German takeover of European countries during WW2) as well as administrative dominance (e.g., Soviet control over satellite states in Eastern Europe and Mongolia) count here.

Demonym: name for residents or natives of a particular place, which is derived from the name of that particular place.

Ethnocide: deliberate and systematic destruction of the culture of an ethnic group (including their language)

Genocide: deliberate killing of a large group of people, especially those of a particular group or nation.