

# SIDE Metadata Codebook

Carl Müller-Crepon and Philipp Hunziker

January 2, 2018

The SIDE metadata described herein is available from: <https://side.ethz.ch>. It includes the following variables:

## **sideid**

Unique identifier for every map layer available. It identifies the data for an ethnic group in a given country, year, and DHS survey. Structure: `sideid = side_v1_mapid_layer`. The `sideid` matches the asc-filename for each map. It can thus be used to bulk-download the SIDE data. Each file is accessible via: <https://side.ethz.ch/raw/v1/ENTER-sideid-HERE.asc>.

## **mapid**

Unique identifier for the ensemble of group-level maps created of one DHS survey in a given country. Each `mapid` is associated with the convex hull of all DHS clusters used for the imputation of the data. Convex hulls can be downloaded from: [https://side.ethz.ch/raw/v1/conv\\_hull/side\\_v1\\_convhull\\_ENTER-mapid-HERE.asc](https://side.ethz.ch/raw/v1/conv_hull/side_v1_convhull_ENTER-mapid-HERE.asc).

## **layer**

Layer-identifier of one group-level map; nested within `mapid`.

## **country**

Name of the country of a group-map.

## **cowcode**

Correlates-of-war code of this country.

## **groupname**

Name of the group as defined by the DHS. "Other" denotes ensemble of groups named "other" in the DHS data, and those that are too small to be effectively interpolated (see paper). If `marker == "ethnic.religion"`, then the `groupname` is a group's ethnic and religious name, separated by a dot.

## **marker**

One of ("ethnic","religion","ethnic.religion"), denoting the type of group.

## **year**

The year of the DHS survey corresponding to a map.

**dhs.round; dhs.subround**

The DHS survey round and subround. Sometimes, multiple surveys were conducted in the same country during one wave of the DHS survey.

**sample.size**

The number of geocoded DHS clusters used for interpolation.

**fit.comb**

The combined leave-one-out-cross-validation fit of the KNN and TPS models used for interpolation this `mapid`. See paper for further details.

**tps.\***

Parameters and fit of the Thin Plate Spline Model used for this `mapid` See paper for further details.

**knn.\***

Parameters and fit of the K-Nearest-Neighbours Model used for this `mapid`. See paper for further details.

**Reference**

Müller-Crepon, Carl & Philipp Hunziker. (2018). New Spatial Data on Ethnicity: Introducing SIDE. *Journal of Peace Research*, forthcoming.