

# Cshapes 2.0: Documentation

Guy Schvitz

[guy.schvitz@gess.ethz.ch](mailto:guy.schvitz@gess.ethz.ch)

December 21, 2018

## 1 Overview

Cshapes 2.0 is a GIS dataset that provides historical maps of independent states and dependent territories from 1886 to 2017. Throughout this period, the data records changes in each country's borders, political status<sup>1</sup>, and other relevant information, such as changes in a country's capital. The final dataset comes with an extensive documentation that contains background information on (1): changes in a country's status, (2) changes in a country's borders and (3) where applicable, coding decisions. This documentation uses the same basic layout as the [EPR-Atlas](#), which provides detailed information on ethnic groups in each country.

## 2 The Dataset

Before turning to the task itself, here is a brief overview of the structure of the dataset. The data is stored as a GIS (vector) dataset, in which countries are represented by polygons, along with attribute information. More specifically, the dataset consists of a table, in which the columns/variables contain relevant attribute information, such as a country's name and political status. An additional column contains a polygon that represents each country's borders on the map. The rows of the table correspond to country-periods, i.e. periods in which a country's borders and attributes remained unchanged. For each change in any of a country's attributes or borders, the table contains a new row, with its own start- and enddates. Your task will be to provide background information on two types of change: (1) Changes in a country's borders,

---

<sup>1</sup>I.e. a change from a status as a colony to an independent state.

and (2) changes in a country's political status. To provide an overview, here are all the variables stored in the Cshapes dataset.

- **gwcode**: Country identifier based on ? for independent states and ? for dependencies.
- **statename**: State name (Aliases in brackets)
- **startdate**: Start date of the country-period
- **enddate**: End date of the country-period
- **status**: Political status of the territory: 1: Independent state. 2: Colony. 3: Protectorate. 4: Leased territory. 5: Occupied.
- **ruledby**: For all dependencies we record the gwcode of the state that holds sovereignty over this territory (e.g. for all british colonies: 200).
- **capname**: Name of the capital.
- **caplong**: Longitude of the capital
- **caplat**: Latitude of the capital
- **area\_sqkm**: Size of the country's territory (sqkm)
- **b\_def**: Dummy variable indicated if a country's borders were clearly defined (1) or partly undefined (0). In many colonies of the late 19th century, large parts of the border remained undefined.
- **fid**: Feature id, i.e. unique id of every row
- **geom\_id**: Unique id of every geometry

### 3 Coding Task

The documentation should provide background information on changes in a country's territory and or its political status in a chronological order. Each entry should be properly referenced and should be relatively short (1-5 sentences in most cases.). In addition, the document should also refer to the map material that was used. A file with all the maps used for each country can be found here: `/polybox/shared/Cshapes Documentation Final/Map Sources/map_sources_list.csv`

We use Latex to create the documentation. Each chapter should be between half a page and maximum three pages long. For each country, I have created an empty Latex template file, which you can edit directly. I have also created a BibTex bibliography file you can use for citations and to which you can add further sources. The files are stored in the polybox-folder<sup>2</sup> and are named as follows: `gwcode.countryname.tex`.

The template file contains the following fields and sections:

- **Country name:** Enter the country name as indicated in the dataset. Put aliases in parentheses (E.g. Ghana (Gold Coast)).
- **Country ID:** Enter the country code, as indicated by the file name.
- **Startdate:** Enter the earliest start-date in our dataset (founding of country or earliest observation in the dataset)
- **Enddate:** Enter the latest end-date in our dataset.
- **Overview:** Where applicable: Briefly provide the most important information on the following points: Is the country a predecessor/successor of another state? (e.g. German Democratic Republic, gwcode 265 was a successor of Germany, gwcode 250). Or did the country cease to exist for a period of time? For example, Lithuania did not exist as a distinct territorial unit between 1945-1991, as it was part of the Soviet Union during this period. If none of this applies, remove the entire section.
- **Changes:** List the changes recorded in our dataset in chronological order, and provide a brief description. First, specify the exact date of the change in YYYY.MM.DD format. In some cases, our dataset does not yet list the correct date, as our source datasets often did not provide exact dates. In these cases, we used the first of January, or the first of the month as a default. If you find that the dates of territorial changes in our dataset are incorrect, please note this in a separate word-file that lists all inconsistencies. Second, briefly describe the event that took place (e.g. territorial gain/loss to country XYZ). Make sure to mention the political events that led to the change, such as treaties, declarations of independence, formal annexations, purchases, and provide adequate sources. Note that we only code *de jure* changes in territory and sovereignty. Therefore, we code the day on which for example a treaty was signed as the day a territorial change occurred.

---

<sup>2</sup>"/polybox/shared/Cshapes Documentation Final/New.tex.files"

- **Coding Decisions:** If our coders provided information on their coding decisions, briefly describe the main coding decisions we made.<sup>3</sup>. Usually coding decisions concern potential instances of territorial change that we did not code as such. Our coding rules excluded territorial changes in the following instances:
  - Territorial changes below 10'000 sqkm
  - Territorial exchanges between states that were unrecognized, and were reversed at some point (E.g. Japan's occupation of Manchuria).
  - Cases in which a *de-facto* secession occurred that lacked international support and recognition (e.g. Biafra).
  - Territorial changes that occurred among warring parties during wartime, but were reversed right after the end of the war (here, we only coded the status quo before and after the war).
  - Cases in which borders remained undefined but no evidence of territorial changes could be found. For example, a border between Saudi Arabia and Oman was largely inexistent before 1952. Prior to this period however, we could not find any evidence that territory was exchanged between the two prior to this period. In these cases, we backdated the current borders to 1886.
  - In some cases, maps of the same region contradict eachother, which was often due to conflicting territorial claims. In these cases, we opted for the map that corresponded most to the region's current borders.
  - In some cases, maps were simply inaccurate, showing slightly different borders compared to today. If we did not find any evidence of territorial changes, we ignored these discrepancies and backdated the current boundaries.
- **Map Sources:** Finally, list the maps that we used as sources to geo-code each country period. We enter these references manually, in order to separate the map entries from the overall bibliography. Follow the citation rules shown in the template file.

The overall procedure is as follows:

1. Start with the first unedited file in the "New\_Tex\_Files" folder.

---

<sup>3</sup>I will also fill in this part myself

2. Check if there is not already a .pdf version of the file in the "Cshapes Documentation Final" folder. If there is one, you can copy the entries from the .pdf version, if these are correct. If there is none, proceed editing the .tex file.
3. The folder "Ctry\_Level\_Change\_Maps" contains a pdf for every country that provides a map of each country-period, along with the main relevant variables. For changes in a country's borders or status, we need an entry under **Changes**, as explained above. Not that not every country-period corresponds to a change that needs to be listed in the documentation, as changes may also stem from changes in the capital city, or other attributes.
4. After listing the changes for each country, cite the maps that were used to code each change.

### 3.1 Notes

Some additional notes:

- Obviously, most border changes affected at least two countries at the same time. In these instances, you can copy your **change** entry for one country, and paste it to the file for the neighboring country if it was affected by the same change, adding perhaps only minor adjustments. This will ensure consistency and will save you some time.
- For any questions of clarification issues, contact Guy.

## References