Step-by-step import guide – contribute a dataset

Initial version of the codebook was created by Phillip Hocks and Jan Fischer (University of Bremen) in 2015. This work was funded by a small research grant of the University of Bremen (M8 Plus).

More information: *www.partyfacts.org – github.com/hdigital/partyfacts*

Introduction

We encourage you to participate and contribute to the Party Facts project by adding party information from your dataset. Through adding your parties you will be able to link your data with well established datasets from the political science community.

In Party Facts, we collect and connect information from social science datasets that cover political parties. We do not include the original dataset but extract party information and add it to Party Facts where it can be linked. Ideally, the list of parties includes a party name in the original language and an English name, some information about the time the party existed and size information (vote share). Having all this information makes adding and linking parties through Party Facts easier.

The import and maintenance of data is done via a public Github repository but the main work of linking parties takes place on the website. Github is a file provider for software and project development, which allows us to coordinate, administrate, archive and adapt the project and datasets perfectly. Github enables you to contribute your list of parties to the project with a simple upload. The public repository folder is available at *github.com/hdigital/partyfacts*, where you can upload your data folder including the required files mentioned below. Afterwards we will continue the process by importing your list of parties into the project database at *www.partyfacts.org*.

The list of parties from your dataset has to follow some guidelines. Before uploading or submitting your dataset, please make sure that the party list is formatted correctly and follows the points we present below.

Your Dataset

Inclusion criteria

Every political science dataset dealing with parties and elections may be added to the Party Facts project.

Format

Dataset file format

The datafile has to be in csv (comma-separated-value) format. If your data is currently not, please convert it to a csv file.

utf-8

Please make sure your file is utf-8 encoded. Every csv file can be encoded in utf-8 (save option) and iit guarantees an error free display of string-variables.

Variables

If your dataset contains additional information (eg. party positions, party leaders etc.) only the respective variables of the Party Facts project will be imported and displayed. You may exclude additional variables in the uploaded file.

Your dataset should include the following information:

- **Country name short** – Country's ISO-code, three digit > abbreviation of country name.

- **Name short** – Common abbreviation of the party name in language > of origin.

- **Name** – Name in language of origin

- **Name English** – Translation of the party name into English

- **Name other** – Party name in languages with non-latin letters

- **Year first** – Foundation of the party
  > If no information is available, the first election occurrence may > be entered

- **Year last** – Dissolvement of the party
  > If no information is available, the last election occurrence may > be entered. Leave empty if the party exists as of today.

- **Share** – Vote share won in a national general election used to > identify the party. Use maximum vote share if information is > available.
  > The decimal place is indicated by a dot and not a comma

- **Share year** – Year of vote share inserted at "share"

Strictly required are *country* and *party name* information.

Your Script

In Party Facts, we import a list of parties extracted from a dataset. Ideally, you provide a script in which the party list is extracted from the dataset including the information needed for a Party Facts import. We prefer R scripts but Stata or other scripts may work as well.

The script helps to understand the origin of the party list and allows us to solve potential problems.

Readme file

To provide information about the dataset it is important to submit a *readme* file together with your dataset. This readme file will be used as an information source on Github and the information displayed on the Party Facts website (*see here*).

For a consistent workflow and information, please make sure your readme file matches the following criteria:

Name: readme.md (markdown)

This is a markdown file which can be either created directly in github or via the texteditor

Template

Please provide the Readme file in the following style.

```
## Source

(...)

## Credits (optional)

(...) name -- institution -- year(s)

## Import

(...)

## Comments (optional)

(...)

## Todo later (optional)

(...)
```

Upload

To import your party list, you need to create a folder with the respective files. The folder has to be named with a single word identifying your data or project.

The folder needs to contain:

- Data file – named with the same word as the folder .csv
- Readme file – *readme.md*
- Script file – extracting party information for Party Facts import

The folder may contain your original dataset. The dataset can be ignored in Github if it includes information required to run the script but should not be published on the Party Facts Github page. In this case, please provide the dataset with the name *source__*filename (two underscores). Files and folders starting with *'source__'* will be ignored in Github and thus not be available publicly.

You can either upload the file directly on the public github repository yourself or send it to one of the project maintainers – Holger Döring or Sven Regel.

Import to www.partyfacts.org

The import into the Party Facts website is done by the Party Facts Team. Afterwards you are ready to link your dataset to the existing core parties and to combine your data with party information from many existing sources.