

Familytreemap User Guide

Dirk Struve

phylofriend at projectory.de

<https://github.com/yogischogi/familytreemap/>

June 25, 2017

Contents

1	Introduction	1
2	Command line options	2
3	Installation	2
3.1	Linux Mint	2
3.2	Windows, FreeBSD, Mac OS X	3
4	First usage	3
5	Calculating relative concentrations	3
6	Does England belong to the United Kingdom?	4
7	How to create a heat map	5

1 Introduction

For the lost ones.

Familytreemap calculates population frequencies (relative or absolute) from Family Tree DNA projects.

It's main purpose is to create intensity maps of project members.

2 Command line options

Familytreemap is a command line program. It is invoked by

`familytreemap <options>`

Options may be given in arbitrary order.

-help Prints available program options.

-totalsin Totals in: Total number of testers from each country.

-in Input file that contains a table with Family Tree DNA project data in CSV format. This is usually the project spreadsheet that lists the STR results.

-col Number of the column that contains the country names.

-out Output file that contains population frequencies in CSV format.

-sumuk (sum UK) If **-sumuk=true** the number of testers from England, Wales, Scotland and Northern Ireland is added to United Kingdom. Default is **-sumuk=false** because this is the way the data is reported by Family Tree DNA.

-statout Filename for elaborate statistical information.

3 Installation

3.1 Linux Mint

1. Make sure that the Go programming language is installed. You can install it by typing
`sudo apt-get install golang`
2. Read the Go [Getting Started](#) guide. Make sure to set your `GOPATH` variable and include it in your `PATH` so that Go programs can be found.
3. Fetch the familytreemap program with
`go get github.com/yogischogi/familytreemap`
4. Install the program with
`go install github.com/yogischogi/familytreemap`

3.2 Windows, FreeBSD, Mac OS X

1. Read the Go [Getting Started](#) guide and install the Go programming language. Make sure to set your *GOPATH* variable and include it in your *PATH* so that Go programs can be found.
2. Fetch the familytreemap program with

```
go get github.com/yogischogi/familytreemap
```
3. Install the program with

```
go install github.com/yogischogi/familytreemap
```

4 First usage

1. Go to a Family Tree DNA project on the web and open the page containing the DNA results in a web browser.
2. Copy the results into a spreadsheet and save it in CSV (Comma Separated Values) format. For this example name it *projectdata.csv*.

If you are an administrator you can directly download the spreadsheet from your Family Tree DNA GAP account.
3. Open a command line interpreter and go to the directory where your files are.
4. Determine the number of the column which contains the countries. Often this is 3, familytreemap's default value.
5. Issue a command to test if the program works:

```
familytreemap -in=projectdata.csv -out=result.csv -col=3
```


If everything works correctly, the file *result.csv* contains a list of countries with the number of testers from each country.

5 Calculating relative concentrations

Until now we have only determined how many persons from each country took a genetic test. This results in a bias because Family Tree DNA usually has much more customers from highly populated Western countries than from the rest of the world.

Often we would like the percentage value of how many persons from a specific country belong to a certain haplogroup. For this we must divide the

number of persons from that haplogroup by the total number of testers from the country.

Family Tree DNA does not provide the total number of testers from each country but it does provide the number of 12-marker testers. This should be a very good approximation because nearly every male customer has tested for at least 12 markers.

To view the number of 12-marker testers you can sign into Family Tree DNA and go to *Ancestral Origins*.

To calculate relative frequencies the Familytreemap program needs a CSV file that contains the total number of testers from each country. You have to create this file yourself. The format is very simple. A basic example looks like this:

```
Country,Testers
Belarus,1000
Belgium,2000
Brazil,100
```

If you save your file as *12-marker-testers.csv* you can start to calculate a relative distribution by using the `-totalsin` option:

```
familytreemap -totalsin=12-marker-testers.csv -in=projectdata.csv
-out=result.csv -col=3
```

The file *results.csv* should now contain the percentage value of persons belonging to a specific project or haplogroup relative to the total number of 12-marker testers from each country.

6 Does England belong to the United Kingdom?

Family Tree DNA reports the number of testers according to self reported origins. They do not apply any post-processing. As a result persons from England, Wales, Scotland and Northern Ireland are not counted as belonging to the United Kingdom.

To get the correct count Familytreemap provides the `-sumuk` (sum UK) option, which adds the testers from England, Wales, Scotland and Northern Ireland to the United Kingdom. Example:

```
familytreemap -totalsin=12-marker-testers.csv -in=projectdata.csv
-out=result.csv -col=3 -sumuk=true
```

7 How to create a heat map

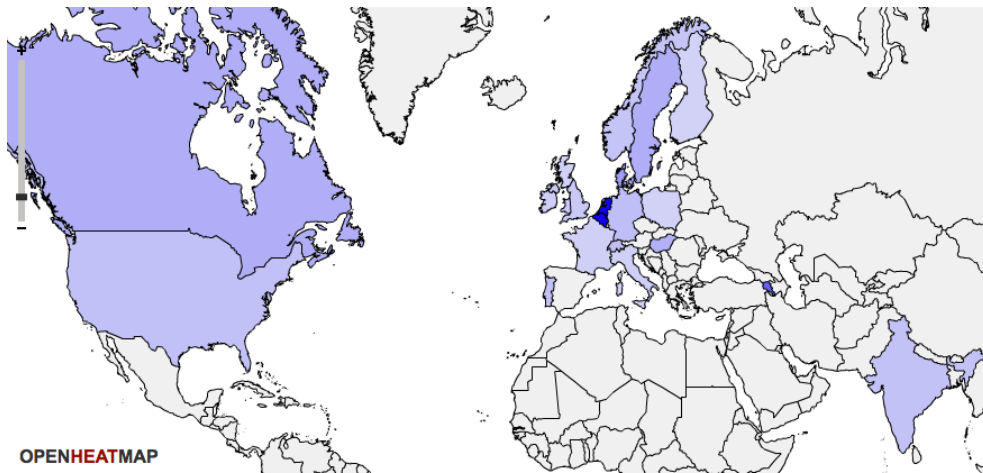


Figure 1: Geographical distribution of persons belonging to a haplogroup. Created with [OpenHeatMap](http://www.openheatmap.com).

1. Run the familytreemap program:

```
familytreemap -totalsin=12-marker-testers.csv -in=projectdata.csv  
-out=result.csv -col=3 -sumuk=true
```
2. Open your web browser and go to <http://www.openheatmap.com>.
 - (a) Click on *Create your map*.
 - (b) *Excel or CSV file*.
 - (c) *Upload* your results file, in this example *result.csv*.
 - (d) *View your map* and adjust the settings until you like it.
 - (e) *Save & view*
3. You are done! You can share your map via social networks or take a screenshot of it.