
med_to_csv

Release 0.9

Jean-Emmanuel Longueville, Myriam Hanna, Nino Mulac, Marcello

Jul 06, 2023

CONTENTS:

- 1 Introduction 1**
 - 1.1 About 1
 - 1.2 Aim of the project 1
 - 1.3 How does it work? 1
- 2 Getting Started 3**
 - 2.1 Requirements 3
 - 2.2 Installation 3
- 3 Command line interface 5**
 - 3.1 How to use CLI 5
 - 3.2 Call example 5
 - 3.3 Details 5
- 4 Graphique User Interface (GUI) 7**
 - 4.1 How to use GUI 7
- 5 Config file yaml 11**
 - 5.1 Set up 11
 - 5.2 Templates 12
- 6 FAQs 15**
- 7 License 17**
- 8 Indices and tables 19**

INTRODUCTION

1.1 About

med_to_csv is a software that converts med associates data files into tidy data.

1.2 Aim of the project

This project aims at facilitating the wrangling and the analysis of data by parsing raw data files obtained from MedAsociates operant chambers, in a given folder, and creating a tidy CSV file containing the information required.

1.3 How does it work?

Since med associates allows to fill freely the information needed, the resulting data layout will differ from user to user. In this software, the user should specify in a **config file** written in YAML:

1. **Which** information to retrieve (i.e naming specific variables) and
2. **Where** this information is present in the original data file (i.e. in which line the variable is written).

The software generates a CSV file in a table format where each variable mentionned by the user produces a column and each observation produces a row.

Note: If arrays are present, unique values are repeated to match the length of longest array.

GETTING STARTED

To get a local copy up and running follow these simple steps

2.1 Requirements

- Python ≥ 3.6

2.2 Installation

Open a terminal and copy the following command lines.

1. Clone the repository

```
git clone https://sourcesup.renater.fr/anonscm/git/medanalysis/medanalysis.git
```

2. Install the required packages

```
pip install numpy argparse pyaml pandas  
pip install Gooey #For graphical interface
```

3. Install software

```
python setup.py install #On gnu-linux /MacOs Systems  
python setup.py install --user #On windows Sytems
```

or

```
pip install .
```


COMMAND LINE INTERFACE

3.1 How to use CLI

You can use `med_to_csv` in CLI or GUI. Here we develop the CLI interface. This interface takes 3 parameters :

- 1st is the path to your directory or file output of Medassociate
- 2nd is the path to the `config file`
- 3rd is the path to the output file which must end by `.csv`

3.2 Call example

Run the following command on the file or directory containing the datas

```
med_to_csv_cli my_results my_config_file.yml output_file.csv  
.. warning:: Your path must contain only ASCII characters (No : éèàï)
```

3.3 Details

GRAPHIQUE USER INTERFACE (GUI)

4.1 How to use GUI

med_to_csv_gui

This software version use Goocy module.

You can find it [here](#)

When you first open the software, the following page opens. The first step is to browse for the [config YAML](#) file that you wrote earlier.

med_to_csv

Help

This software converts med associates data files into a unique csv file.

For more information about the config file please check the README or documentation.
You must indicate a path or a file with your data otherwise the software will return an error.

Required Arguments

option
Path to option config file.
C:\Users\jean\datas\config.yml

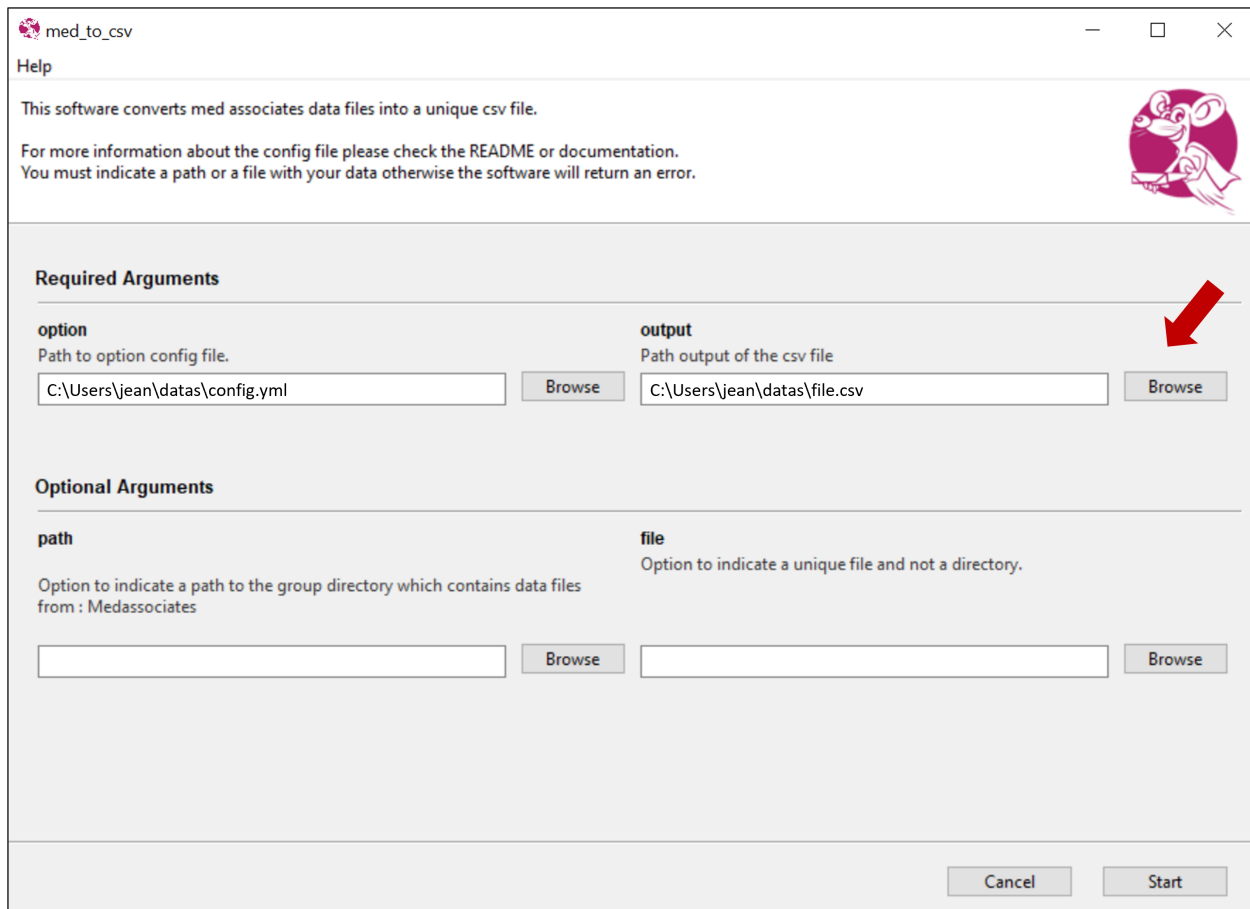
output
Path output of the csv file

Optional Arguments

path
Option to indicate a path to the group directory which contains data files
from : Medassociates

file
Option to indicate a unique file and not a directory.

Then, you should mention the output CSV file (Path to the location and name of the output file)



med_to_csv

Help

This software converts med associates data files into a unique csv file.

For more information about the config file please check the README or documentation.
You must indicate a path or a file with your data otherwise the software will return an error.

Required Arguments

option Path to option config file.	output Path output of the csv file
<input type="text" value="C:\Users\jean\datas\config.yml"/> <input type="button" value="Browse"/>	<input type="text" value="C:\Users\jean\datas\file.csv"/> <input type="button" value="Browse"/>

Optional Arguments

path Option to indicate a path to the group directory which contains data files from : Medassociates	file Option to indicate a unique file and not a directory.
<input type="text"/> <input type="button" value="Browse"/>	<input type="text"/> <input type="button" value="Browse"/>

After that you have two options. You either open the **folder** (on the left box) that contains the med associates datas (make sure to mention the main folder that contains the different folders containing the raw data) OR you enter the **file** (on the right box) if all your data is in one file.

med_to_csv

Help

This software converts med associates data files into a unique csv file.

For more information about the config file please check the README or documentation.
You must indicate a path or a file with your data otherwise the software will return an error.

Required Arguments

option
Path to option config file.
C:\Users\jean\datas\config.yml Browse

output
Path output of the csv file
C:\Users\jean\datas\file.csv Browse

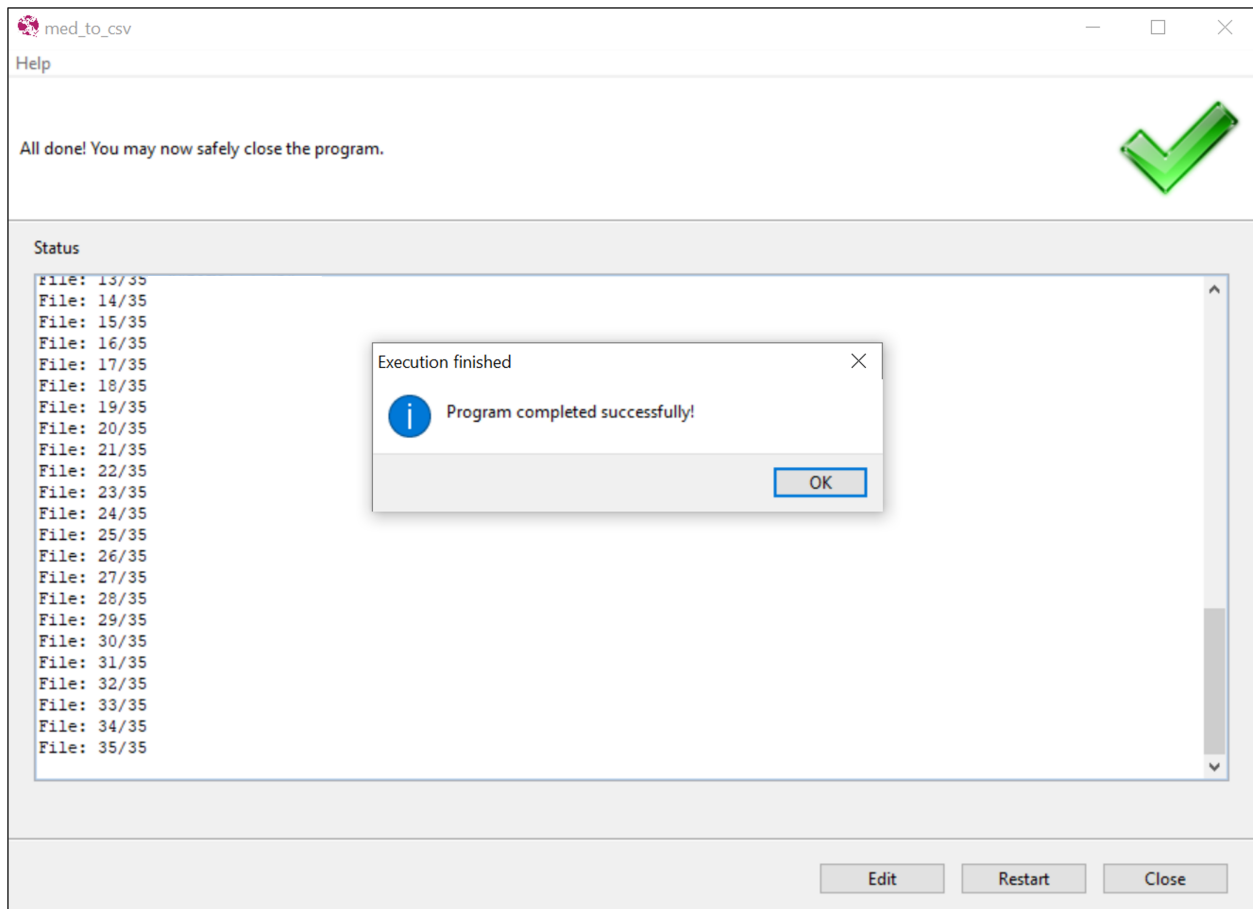
Optional Arguments

path
Option to indicate a path to the group directory which contains data files
from : Medassociates
C:\Users\jean\datas Browse

file
Option to indicate a unique file and not a directory.
Browse

Cancel Start

Then click on **start** and watch the software run.



CONFIG FILE YAML

5.1 Set up

The config file is a necessary file that gives the setup of your data for the software. You can find an example in the config.yml file which contains all possibilities of setup. Med-PC can create different types of files:

1. One-column file, where the data is represented in only one column, and
2. Annotated file, where the data is labelled by the letters from the Med-Associate script

Note: You can mix directories of one-column files and directories of annotated files in one folder.

1. There are three types of parameters:

- info_col : One column file
- info_lab : labelled file
- options : further options

Important: Either info_col or info_lab are required, if your data contains a mix of one-column files and labelled files you can mention both parameters

2. For the first two types of parameters (info_col or info_lab):

You need to indicate the information in this format: “Key : Value”

- Keys are the column names that you want in the output file
- Values are:
 - for info_col: row number - 1 (Python starts index at 0)
 - for info_lab: letters used in medAssociate exercise

3. Options:

- remove_zero_ending : True or False to keep or remove Zeros at the end of arrays. In the Med-PC language, the number of slots in an array is defined by the user. If the number of value is < number of slot, then MedPC will add 0 for missing values. If you do not want these data in your csv file, specify the option True in your yaml file.
- filter : a string present in the directories name present in the file directories name that you want to use to select a subgroup of datafiles or directories. For example, you can name test session with the string “test” and use it to filter data excluding data that does not contain the string “test” in their name.

- cut : Some users “tag” data in an array in “TimeStamp.CodeEvent”. For example, they can add a decimal to the time stamp of the event to indicate a specific condition. These data can be retrieved in two separate columns using this option. The value must be a list of list of 4 elements :
 - key to cut
 - separator: usually the dot character
 - column name of first part (eg: TimeStamp)
 - column name of second part (eg: CodeEvent)
 - eval : for some columns the values must be the result of a Python command line (e.g to get information in a path). This must be a python dictionary with Key (column name) and value (short command in a string format).

4. The following Keys are med associate keywords only usable with labelled file :

- Start Date (automatically added)
- End Date
- Subject
- Experiment
- Group
- Box
- Start Time (automatically added)
- End Time
- MSN: Medassociate exercice names

Note: Templates are available in the config.yml file or see below

5.2 Templates

```
infos_col:
  subject: 6
  box: 9
  delay: 66
  step_at_endsess: 73
  sk_dur_at_endsess: 88
  reinforced_resp: 90
  reinforcers: 91
  res_during_reward_deliv: 92
  to_resp: 93
  delivery_s: 94
  to_s: 95
  shock_s: 96
  resets_number: 97
  reinforcers_max_sk_dur: 98
  mn_left_non_resp: 99
  resp_cur_sk_dur: 100
```

(continues on next page)

(continued from previous page)

```
final_sk_dur: 101
base_value: [120, "end"]

infos_lab:
  subject: "Subject"
  box: "Box"
  delay: "B"
  step_at_endsess: "I"
  sk_dur_at_endsess: "X"
  reinforced_resp: ["Z", 0]
  reinforcers: ["Z", 1]
  res_during_reward_deliv: ["Z", 2]
  to_resp: ["Z", 3]
  delivery_s: ["Z", 4]
  to_s: ["Z", 5]
  shock_s: ["Z", 6]
  resets_number: ["Z", 7]
  reinforcers_max_sk_dur: ["Z", 8]
  mn_left_non_resp: ["Z", 9]
  resp_cur_sk_dur: ["Z", 10]
  final_sk_dur: ["Z", 11]
  base_value: ["Z", 30, "end"]

options:
  remove_zero_ending: True #Or False
  #https://docs.python.org/3/howto/regex.html
  filter: "shock"
  cut : ["base_value", ".", "time_stamp", "shock_step"]
  eval: {"prog": "path_file.split('/')[-1]"}
```


LICENSE

This software is distributed on GPLv3.0 <<https://www.gnu.org/licenses/gpl-3.0.en.html>>_ License.

You can use it, distributed it, and naturally read the code. Please if you use and like it add a star on our github <https://github.com/hedjour/med_to_csv>_ page.

INDICES AND TABLES

- `genindex`
- `modindex`
- `search`