# MIP Quality Control Tool Quick Guide (for version 0.1.0)

## mipqctool as python script

### Requirements

* Python 3 with pip
* Installed version of Perl (needed for pdf export)
* Installed LaTex compiler (needed for pdf export)

For Debian based distros we run:

$ sudo apt-get update

$ sudo apt-get install python3 python3-pip latexmk texlive-latex-extra

For windows systems the LaTex Compiler can be downloaded from here: <https://miktex.org/download>

And the Perl from here: <https://www.perl.org/get.html>

### Installation

[git clone https://github.com/aueb-wim/DataQualityControlTool.git](https://github.com/aueb-wim/DataQualityControlTool/tree/master/qctool)

$ cd DataQualityControlTool

$ sh install.sh

For Windows OS in command prompt we run in the folder where the setup.py is located the command:

pip install -e .

### Usage

**CLI mode**

For csv dataset we run the command:

qctool -m csv --input\_csv *[dataset csv path]* --meta\_csv *[metadata csv path]* --col\_val *[metadata column name for variable codes/names]* --col\_type *[metadata column name for variable types] (--readable) (--pdf)*

“-m” or “--mode” can take two flags “csv” or “dicom”. Here we use “csv” because the dataset is in csv format.

“--readable” is a flag if we want the reports csv files to have more descriptive column names.

“col\_val“ and “col\_type” are obligatory. They are referred to columns names of the metadata csv. The “col\_val” is the name of the column that contains the variables codes used as columns in the datatset csv and the “col\_type” is the name of the column in metadata csv that contains the variables types.

At the moment the tool needs the metadata file to detect the nominal variables. So, the “col\_type” column must be filled with the value “nominal” (is not case sensitive) for the categorical variables in order to work properly. Other types like “int”, “float”, “text”, “numerical”, “date” are not taken into account at the moment.

After the execution, three files will be produced:

* a csv file <dataset\_file> + ‘\_dataset\_report.csv’ containing the Statistical Report of the given dataset.
* A csv file <dataset\_file> + ‘\_report.csv’ containing the Statistical Reports of the variables of the given dataset.
* A pdf or LaTex file <dataset\_file>+’\_report’ containg the above two reports in a readable format.

For a DICOM dataset we run:

qctool -m dicom --root\_folder [folder with dicoms] –report\_xls [path/to/report.xls]

“-m” or “--mode” can take two flags “csv” or “dicom”. Here we use “dicom” because we have a DICOM dataset.

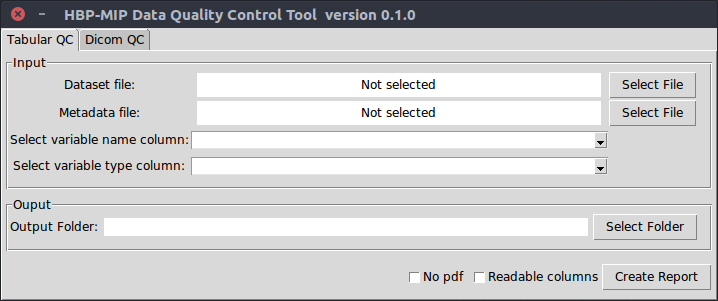
The report is in excel format and contains the header information from all DICOM (dcm) files.

### Gui mode

We run in terminal:

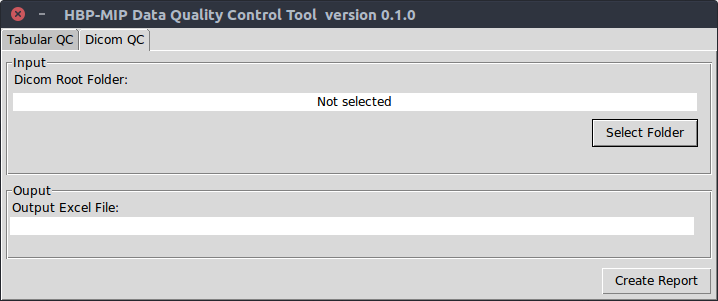
$ qctoolgui

**Tabular(csv) dataset TAB**

For a csv dataset we follow the below steps:

1. Select the dataset csv file
2. Select the metadata csv file
3. Select the column name which contains the variable name in metadata file (similar to col\_var argument in cli mode)
4. Select the column name which contains the variable type in metadata file (silimar to col\_type argument in cli mode)
5. Select the output folder where the 3 reports files will be saved
6. Check the box “Readable columns” if we want bigger and more descriptive columns in reports in csv format.
7. Check the box “No pdf” in case we don’t have a latex compiler installed. In that case the tool is produce a .tex file which can be compiled in another machine or using an on-line LaTex compiler.
8. Click “Create Report”. An info message will pop up when the reports

**DICOM dataset TAB**



For a DICOM dataset we follow the below steps:

1. We select the root folder where the subfolders with DICOM files(.dcm) are located.
2. We press “Create Report” button and define the name of the report file (Excel format)

## mipqctool as standalone windows executable

### Requirements

* 64bit Windows OS
* Installed version of Perl (see previous section)
* Installed LaTex compiler (see previous section)

### Installation

Download the windows executable winqc.exe (cli mode) and winqcgui.exe (gui mode) from:

<https://github.com/aueb-wim/DataQualityControlTool/tree/master/windows_executables>

Optionally, we can add the folder where the qctab.exe is located to the system’s PATH variable in order to run the application globally.

### Usage

In cli mode for csv dataset, we run in the command prompt:

*winqc -m csv --input\_csv [dataset csv path] --meta\_csv [metadata csv path] --col\_val [metadata column name for variable codes/names] --col\_type [metadata column name for variable types] (--readable) (--pdf)*

The arguments are the same with the ones that are described above (python script execution).

For DICOM dataset, we run:

winqc -m dicom --root\_folder [folder with dicoms] –report\_xls [path/to/report.xls]

For gui mode we just execute the winqcgui.exe and follow the steps described in the previous section.

Notes: This version of winqc.exe and winqcgui.exe is tested in Windows 7 and Windows 10