

Data Plotting Tool v22

Developed by Suresh Pokhrel - ICIMOD

This tool was developed in Python 3.9

Date of release: July 23, 2022

*** Disclaimer:** THE SOFTWARE IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT. IN NO EVENT SHALL THE AUTHORS OR COPYRIGHT HOLDERS BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER LIABILITY, WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE, ARISING FROM, OUT OF OR IN CONNECTION WITH THE SOFTWARE OR THE USE OR OTHER DEALINGS IN THE SOFTWARE.

Please read the instructions carefully.

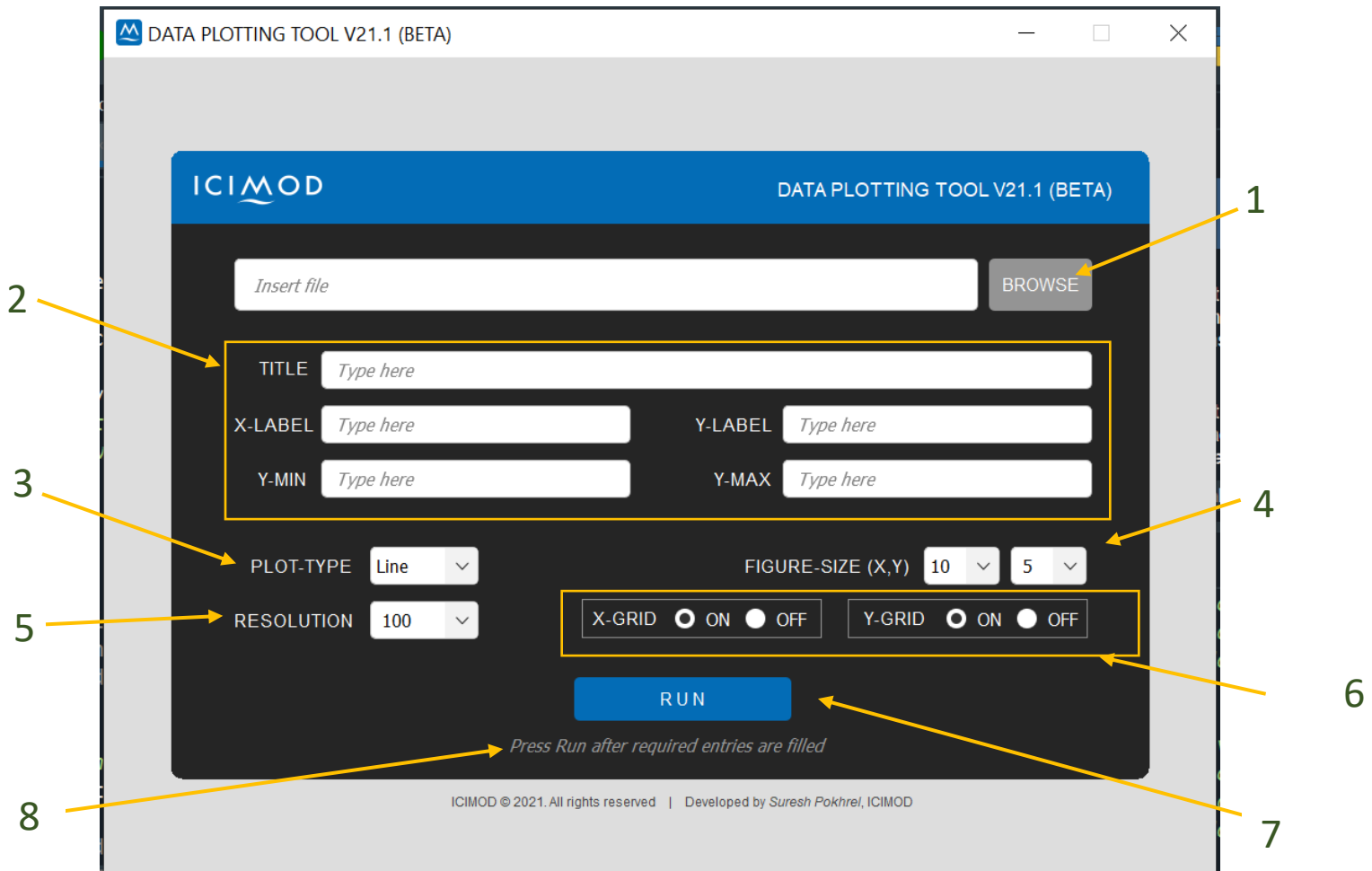
Requirements:

- 1) Runs on Windows 64-bit version only.
- 2) The input file should be in csv format. No other format supported.
- 3) The requirement of the column types of the input file depends upon Plot-Type:
 - For **Line plot** and **Point plot**: The first column should contain common standard of the Date & Time format (i.e. mm/dd/yyyy hh:MM, E.g., October 15, 2020, 1:30 PM should be represented as "10/15/2020 13:30"). Rest other columns can be data columns.
 - For **Bar plot**: The first column should be indexing data and rest other columns can be data columns.
 - For **Box plot**: The first column can be any indexing data which is ignored. Rest other columns can be data columns.
 - For **Scatter plot**: The first column can be any indexing data which is ignored. The second column should be X variables and third column should be Y variables. Rest other columns are ignored.
 - For **Histogram plot**: The first column can be any indexing data which is ignored. Rest other columns can be data columns.
 - For **WindRose plot**: The first column can be any indexing data which is ignored. The second column should be Wind Direction (unit should be in degree) and third column should be Wind Speed. Rest other columns are ignored.
- 4) The provided folder "img" and the files inside this folder should never be renamed, deleted, or relocated. Any alteration on those files and the folder will result in malfunctioning of the tool.

Procedures:

Run the exe file "Data_plotting_tool_22".

The software window should appear after few seconds which looks like a window image provided below:



Now user should provide inputs. The details are:

- 1) Press "BROWSE" button to locate the input file and the file path will be displayed on the left side of the button.
- 2) The information of the entries to be filled as per the options are:
TITLE = Title of the graph (appears on the Top-Centre of the graph)
X-LABEL = The name to be provided for X-axis label (e.g., Date). Ignore this for WindRose plot.

Y-LABEL = The name to be provided for Y-axis label (e.g., PM10, ug/m3)

Y-MIN = The minimum value of Y for which a user wants the graph to be started from. Ignore this for WindRose plot.

Y-MAX = The maximum value of Y for which a user wants the graph to be ended. Ignore this for WindRose plot.

** All these fields described in point 2 are optional one. User can leave blank if not required. The subscript and superscript used for the basic label names or title names are handled by the tool automatically. The user can refer the table below to provide the appropriate input text for desired output text.

User Input	Converted Output	Remarks
PM10	PM ₁₀	PM ₁₀
PM2.5	PM _{2.5}	PM _{2.5}
PM1	PM ₁	PM ₁
O3	O ₃	O ₃
CO2	CO ₂	CO ₂
SO2	SO ₂	SO ₂
NOx	NO _x	NO _x
NO2	NO ₂	NO ₂
ug/m3	μg/m ³	μg/m ³
DEG_	°	Degree Sign
MICRO_	μ	Micro or Meu Sign

** Note: User Input are case sensitive.

**If user wants customization to subscript or Superscript a text/character then use syntax:

$\$_{xx}$ for subscript and $\xx for superscript (e.g. For m/s², the text should be m/s²)

- 3) User can select the types of plots to be made from "PLOT-TYPE" menu. The available types are Line, Point, Box, Bar, Scatter, Histogram and WindRose.

**If timestamp is/are missing for the Line plot then the straight line will be appeared for those missing periods.

- 4) The size of the figure can be adjusted using "FIGURE-SIZE (X, Y)" menu. The first drop-down option is for length of X-axis and the second is for Y-axis. The unit is in inch. Ignore this for WindRose plot.

- 5) The resolution of the output picture file can be selected using "RESOLUTION" menu. The range is from 100 to 3000 and unit is dot per inch (dpi).
- 6) User can select the "X-GRID" and "Y-GRID" separately with "ON" or "OFF" button. Selecting "ON" will cause grid line to be appeared on its respective axis. Ignore this for WindRose plot.
- 7) Once the user selects all the required options then "RUN" button can be clicked to run program generating output.
- 8) User can view each process taking place or known errors if any, just below the "RUN" button after pressing the "RUN" button.

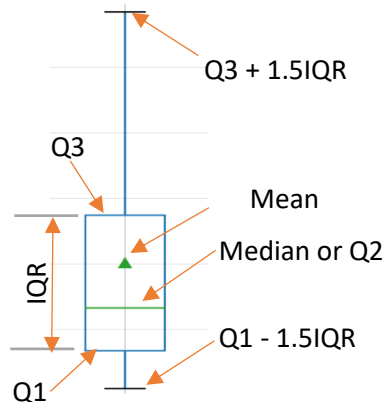
The time taken to process a data depends upon the speed of the PC and the size of the input file.

- 9) If the process is successful, then the Message "Successfully Completed" will be displayed.

Then a beep sound is produced with the opening of the folder containing the output file. The output file path will be the same as of the input file path.

Output:

- The output filename will be same as the input filename but contains pre-fix of timestamp on which the output file is created, to avoid over-writes of existing files.
- If the data contains multiple data columns, then all those data will be plotted in the same axis.
- The description of the boxplot if selected "Box" in the plot type is:



- The small green triangle denotes the sample mean.

- The horizontal green line inside the box indicates the sample median (Q2).
- The lower boundary of the Box is the first quartile (Q1) of the sample whereas the upper boundary (Q3) is the third quartile. Hence, the height of the box represents Interquartile Range (IQR).
- The lower cap of the whisker is the minimum value of the data set and calculated as $Q1 - 1.5IQR$.
- The upper cap of the whisker is the maximum value of the data set and calculated as $Q3 + 1.5IQR$.

For more information, Refer: <http://web.pdx.edu/~stipakb/download/PA551/boxplot.html>

- The description of WindRose plot can be referred from :
https://www.epa.gov/sites/default/files/2019-01/documents/how_to_read_a_wind_rose.pdf
- For description of Histogram plot, refer:
<https://www.statology.org/describe-shape-of-histogram/>

Limitations:

- ❖ It can only plot in a single 2D axis. Subplots are not available.
- ❖ User cannot define the colors of the plot as the multiple columns might be present in the data file. So, defining colors to each of the multiple columns was not found to be feasible. The tool assigns default colors to each of them.
- ❖ The position and font-parameters of title, labels, legends, etc. cannot be configured by the users. All those will be having default values.

Note: If the tool hangs-up for a very long-time during processing without the message of successful completion then it is probably due to system constraints of the PC (usually the RAM) or may be bad input file format. So, for this, user is requested to terminate all the running applications in the PC and try to re-run the tool or check the input file properly.

We would highly appreciate if the developer of the software is acknowledged/mentioned whenever and wherever the software is used to create any output.

*****For further queries or feedbacks, contact at suresh.pokhrel@icimod.org or sureshpokharel91@gmail.com *****