

## Excel macros to more easily customize markers and lines on graphs for clear and accurate visual data presentations

Noritaka Mamorita,<sup>1</sup> Naoya Arisaka,<sup>1</sup> Hidenori Inaoka,<sup>1</sup> Akihiro Takeuchi<sup>2</sup>

<sup>1</sup>Department of Medical Informatics, Kitasato University School of Allied Health Sciences

<sup>2</sup>Department of Medical Informatics, Kitasato University School of Medicine

**Objective:** Parts of Excel graphs, e.g., markers, and lines, can be customized by pointing and modifying lines or markers in small sample sizes; however, it is more difficult to manually customize multiple targets in large sample sizes. We developed macros that more easily customize markers and lines with colors and sizes.

**Methods:** Our Excel macros were developed and tested with Excel 2013 on Windows 10. Macros customize special markers and lines in Excel graphs according to color and width properties that those marker sizes or line widths were set in the right side cells of the data series on an Excel sheet.

**Results:** We coupled original data and property data on each row line with certain key words on an Excel sheet. The macros are: Cell\_color\_to\_RGBhex, Line\_coloring, Line\_width, Marker\_coloring, Marker\_size, etc. We showed that the graph of virtual sample data, 73 series with 4 paired tests, can be customized using macros.

**Conclusions:** Although there were relatively few core source codes of macros for simplicity, in this sample run, those macros easily customized markers and lines of sample virtual data to be colored and sized on an Excel default graph.

**Key words:** Data visualization, Excel graph, Excel macro

### Introduction

Data presentation in figures for scientific publications is the foundation of our collective scientific knowledge and critically important toward further developments.<sup>1</sup> However, authors generally use figures to present summary statistics, instead of providing detailed information about the distribution of the data or showing the full data.<sup>1</sup> Many results in research studies are usually summarized as the mean and standard error or standard deviation in decades.<sup>2</sup> The summarized data often cause readers to incorrectly infer that the data are normally distributed with no outliers. Weissgerber et al.<sup>1</sup> clearly demonstrated that many different data distributions can lead to very similar mean and line graphs. It might be necessary to take a careful look at some new, unknown issues, or error investigations in research and development in various fields.

"The simple graph has brought more information to the data analyst's mind than any other device," Wickham

et al.<sup>3</sup> Weissgerber et al.<sup>1</sup> argued that scatterplots allow readers to quickly assess the direction, magnitude, and distribution of the changes. Some graphing applications are well designed, make graphing easier, and provide additional functionality (e.g., GraphPad Prism, and SigmaPlot).<sup>4</sup> However, these commercial applications may be more expensive than Microsoft Excel. Articles have been published with update task analyses for users with Excel 2007 or 2010.<sup>5,6</sup> Excel is very useful as a database and graphing tool. Additional functions are developed and customized with Excel macro and VBA (Visual Basic for Applications) to answer various user problems. Weissgerber et al.<sup>1,2,7</sup> created and distributed the Excel templates and free web-based tool for presenting continuous data in small sample size studies.

Although parts of Excel graphs, such as axes, labels, markers, and lines, can be customized by researchers using small sample sizes, it is often exceedingly difficult to manually customize many markers and lines in large sample sizes. When designs of markers and lines are not

Received 3 July 2019, accepted 19 July 2019

Correspondence to: Noritaka Mamorita, Department of Medical Informatics, Kitasato University School of Allied Health Sciences 1-15-1 Kitasato, Minami-ku, Sagami-hara, Kanagawa 252-0373, Japan  
E-mail: mamorita@kitasato-u.ac.jp

suitable to present the data by any of the default settings, the user should change all the target objects by manually clicking an object and modifying each design. The larger the sample size, the more difficult it is to display the data in graph form. Therefore, we developed some new Excel macros to more easily customize special markers and lines with colors and widths that accurately display data series in Excel graphs.

### Methods

Our Excel macros below were ordinarily developed and tested with the development tool, Excel 2013 on Windows 10. Macros presented in this document customize special markers and lines in Excel graph according to color properties and width properties that those marker sizes or line widths were set in the right side cells of the data

ID	Test 1	Test 2	Test 3	Test 4	Marker_coloring	Marker_size	Line_width
1	95	95	94	96	FF0000	15	5
2	77	79	75	80	00AA00	20	5
3	69	71	67	75	00FFFF	3	1
4	56	73	83	75	00FFFF	3	1

Line_coloring	Line_segment	Line_segment	Line_segment
00FFFF	FF9900	FF9900	00FFFF
FF9900	00FFFF	FF9900	00FFFF
FF9900	00FFFF	FF9900	00FFFF
FF9900	00FFFF	00FFFF	FF9900

Figure 1. Marker and line properties defined on a data sheet

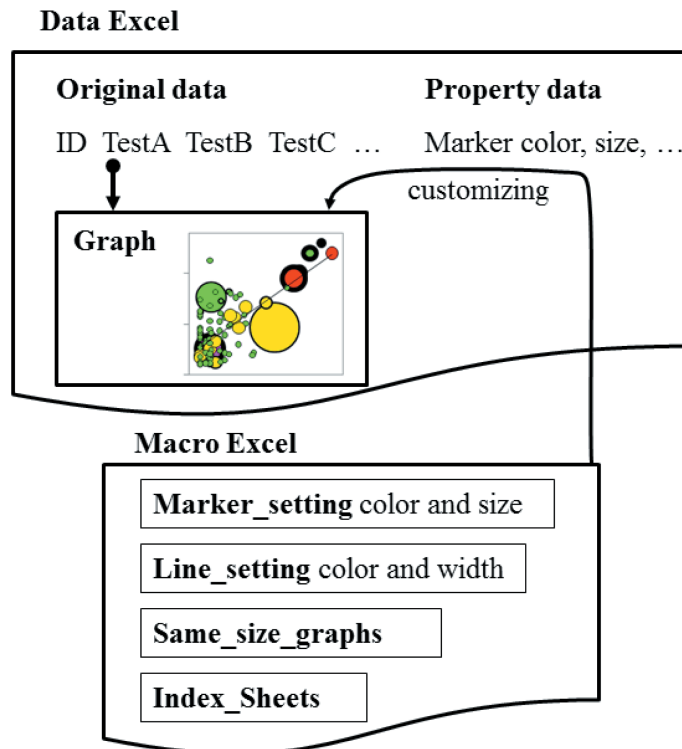


Figure 2. User interface of Excel macros

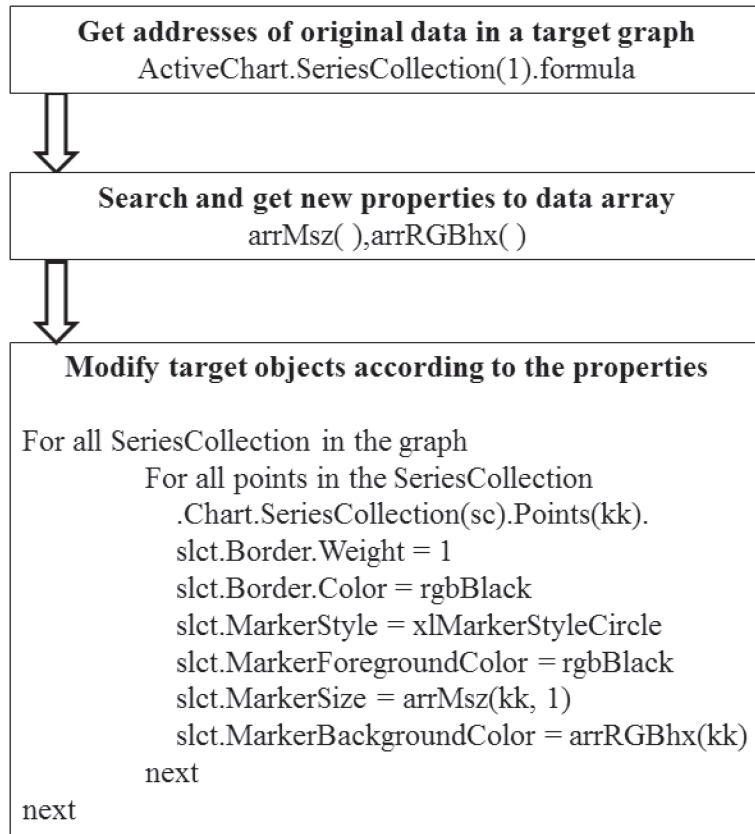


Figure 3. The macro flow chart includes several core source codes.

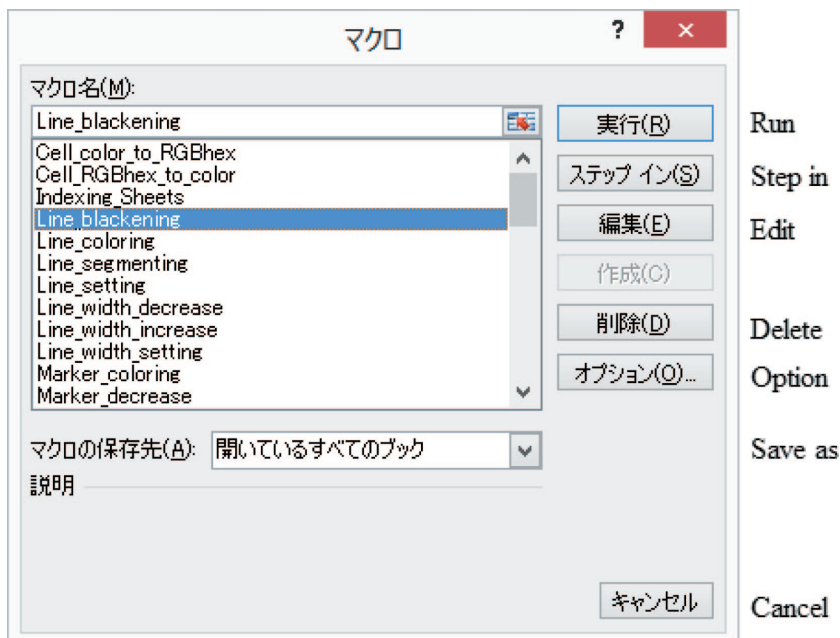


Figure 4. Macro selection window with macro titles listed alphabetically

series in an Excel sheet (Figure 1). The top cell of each property column had a predefined key word, such as: "Marker\_coloring," "Marker\_size," "Line\_width," "Line\_coloring," and "Line\_segment." Red, green, blue (RGB) values were described in hexadecimal numbers in columns of "Marker-coloring" or "Line-coloring." While the values of the "Line\_width" must be greater than 0, the values of the "Marker\_size" must be greater than 1.

When a graph is meant as a target graph, and a special macro is selected via the macro selection window, the macro: (1) gets an address (row and column numbers) from the original data of the target graph to find the property columns near the original data; (2) searches and gets the property data with the key words: "Line\_coloring," "Line\_width," "Marker\_coloring," etc. on the top line of the address sequentially; and (3) modifies target objects on the graph according to the properties (Figure 2). Figure 3 is a flow chart of the macros. The boxes include only a few core source codes for simplicity. The marker line color is set by "Series.MarkerForegroundColor," and the line color of the connecting lines is set by "Series.Format.Line.ForeColor." In the flow chart, the marker style was defined as a circle by "xlMarkerStyleCircle." That was the result of our pros and cons examinations because closed circles were clearer and more suitable in a scatter plot and a series plot than other markers: triangles, squares, diamonds, etc.

## Results

All of the newly created macros are included in the Excel file, "IRmacrosEngl.xlsm (700 kB)." To adapt the macros

to a user's Excel graph, the user should: (1) open (run) both the user's Excel graph and the macro file concurrently (Figure 2), (2) point to a target graph in the user's Excel sheet, (3) open the macro selection window (Figure 4), and (4) select a macro listed in the window.

Each macro title was created describing its function listed in the macro selection window (Figure 4). The macros are listed alphabetically: "Line\_blackening," "Line\_coloring," "Line\_segmenting," "Line\_width\_decrease," "Line\_width\_increase," "Line\_setting," "Line\_width\_setting," "Marker\_coloring," "Marker\_decrease," and "Marker\_increase," etc. "Line\_setting" is a combination of "Line\_coloring" and "Line\_width\_setting." "Line\_segmenting" colors line segments separately. Other related macros were: "Cell\_color\_to\_RGBhex," "Cell\_to\_RGBhex\_color," "Indexing\_Sheets," etc. "Cell\_color\_to\_RGBhex" writes the hexadecimal value of the cell background color to the cell. "Cell\_RGBhex\_to\_color" inversely colors the cell background according to the RGB hexadecimal value defined in each cell. Those macros were used to color the cells displayed in Figure 1. Using the "Same\_size\_graph" makes the size of all the other graphs in that set the same size as the target graph in a sheet at once.

Virtual sample data, 73 series with 4 paired tests, are shown in Figure 1. The original graph is usually created by the Excel default setting (series labels are deleted here) (Figure 5). Although the lines and markers in Figure 5 were designed with the default colors, the marker size, and the line width, they were not clear or suitable to display.

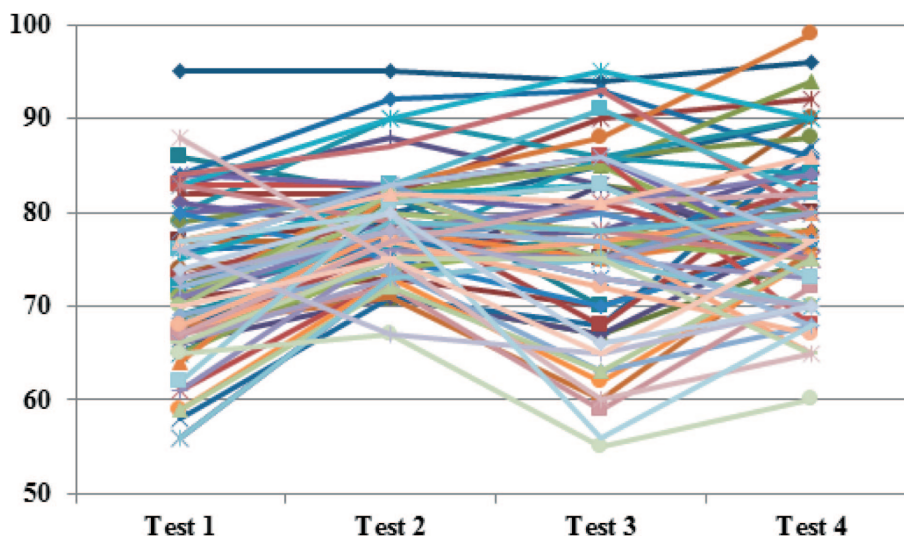


Figure 5. An original graph created by Excel default settings

The lines and markers in Figure 5 were sequentially customized with the macros, "Line\_blakening," "Marker\_sizing," and "Marker\_coloring" as seen in Figure 6A. All the lines in the figures are black. The top line showed the first series in the sheet in Figure 1. Four large red markers were compatible to "20" in the "Marker\_size" column and "FF0000" in the "Marker\_coloring" column. The other small purple and cyan markers were customized in the "Marker\_size" and "Marker\_coloring" columns. The macro "Line\_coloring" sets all the data lines in a figure to blue lines (Figure 6B). The macro "Line\_segmenting" sets each line color to RGBhex defined in the "Line\_segment" column (Figure 6C). The colors of the lines and their meanings are arbitrarily defined by the user as, "increasing between neighboring values" to "00FFFF" and "decreasing values" to "FF9900," e.g., those values were defined by the Excel equation: "=if(C2 < D2, "00FFFF", "FF9900)" (Figure 1)."

The macros "Line\_decrease/increase" and "Marker\_decrease/increase" decrease or increase in a step-wise fashion for all line widths and/or all marker sizes in a graph. All the macros were also adapted to a scatter plot graph as shown in Figure 2.

## Discussion

To discover new trends or findings in the raw data of experiments, it is important to carefully read the raw data and distinguish special data from others. To make distinctive data or trends stand out in Excel graphs, it may be necessary to manually exaggerate markers and lines, not jam data. For all target markers, those manual steps are as follows: pointing to a target marker or line, clicking on the right button of the mouse or pad, and

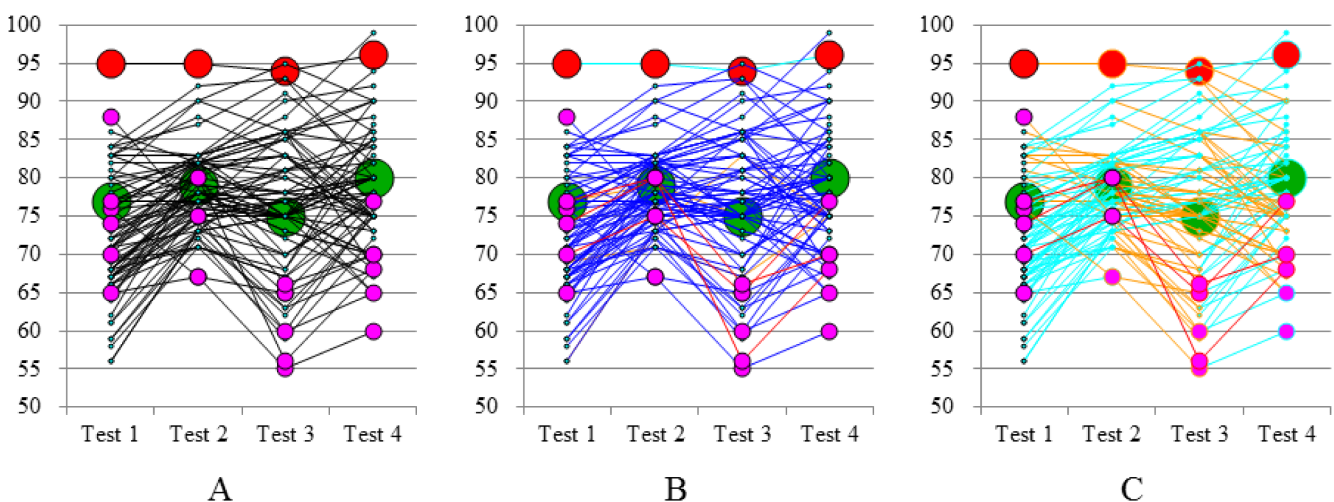
checking option boxes in a property window. However, in large sample sizes, it is difficult to point manually to a target object on many overlapping lines in a graph such as that in Figure 5. Although a macro selection window must be opened in Excel, these newly developed macros can customize target objects without having to manually click on objects.

To do this, we coupled original data and property data on each row line with key words on an Excel sheet. Then we developed the Excel macros to more easily customize special objects with colors and sizes in graphs. We could not find any articles describing this concept or macros in the PubMed database using the key words: "macro, graph, marker, and line." For paid and free graphing software (SPSS, DeltaGraph, R, etc.), it is possible to customize lines and markers with colors and sizes individually with complicated original programming codes. However, our Excel macros can automatically customize target objects by the Excel equation, depending on data.

Our macros are considered to be useful for displaying and observing time series data such as clinical research with larger sample sizes. However, the color of lines and markers needs to specify the red, blue, and green components with hexadecimal values.

Although concrete data of this experiment are not shown here, our newly developed macros were surprisingly easy to use and produced clear and accurate graphs. These types of macros will be further developed to customize any and all objects possible in your Excel graphs.

**Conflicts of Interest:** None



**Figure 6.** The macros modified the original graph (Figure 5) to setting marker and lines (A), to blue lines (B), and to color trend lines (C).

## References

1. Weissgerber TL, Milic NM, Winham SJ, et al. Beyond bar and line graphs: time for a new data presentation paradigm. *PLoS Biol* 2015; 13: e1002128.
2. Weissgerber TL, Garovic VD, Savic M, et al. From static to interactive: transforming data visualization to improve transparency. *PLoS Biol* 2016; 14: e1002484.
3. Wickham H, Grolemund G. *R for Data Science: Import, Tidy, Transform, Visualize, and Model Data*. California: O'Reilly Media, Inc.; 2017.
4. Basics of graphs. Available at: <https://www.linkedin.com/pulse/basics-graphs-jatin-raina>. Accessed May 21, 2019.
5. Dixon MR, Jackson JW, Small SL, et al. Creating single-subject design graphs in Microsoft Excel 2007. *J Appl Behav Anal* 2009; 42, 277-93.
6. Vanselow NR, Bourret JC. Online interactive tutorials for creating graphs with Excel 2007 or 2010. *Behav Anal Pract* 2012; 5: 40-6.
7. Weissgerber TL, Savic M, Winham SJ, et al. Data visualization, bar naked: A free tool for creating interactive graphics. *J Biol Chem* 2017; 292: 20592-8.