1. **Performance comparison among:R, Matlab, esProc, SAS, SPSS, Excel,** *(0 messages)*

   - *Posted by:* David li
   - *Posted on:* August 15 2013 04:46 EDT

The interactive analysis is a circular analysis procedure comprising assumption, verification, and calibration by the analyst to achieve the fuzzy computation goal. For specifics and details, please refer to another article I composed: *Interactive Analysis and Related Tools*.

Because there are so many tools for interactive analysis that not everyone can know all of them, I will only focus on and compare the 7 commonest tools: R, Matlab, esProc, SAS, SPSS, Excel, and SQL. These tools and languages are quite distinctive. In fact, I think there are also quite a few good tools, such as BMDP, Eview, Stata, S-Plus, Octave, Scilab, Mathematica, and minitab, just to name a few. We will discuss them in other essays.

Since it is the comparison on the interactive analysis tools, the project to compare must be also typical and pertinent correspondingly. I listed 5 metrics, and each of them is rated by 5 ★ at the highest. The more ★, the more advantages this tool has regarding this metric.

For example,

UI friendliness: The more ★, the friendlier interface and the easier operations will be.

Technical competence requirement: The more ★, the lower requirements on the technical background like the mathematic algorithm and programming skills.

Stepwise computation: The more ★, the easier to decompose and solve the complex problem.

Support for structured data: The more ★, the easier to carry out the analysis on the structured data.

Fixed algorithm: The more ★, the greater number of fixed algorithms and the stronger functionality are available.

These 5 metrics surely cannot present the ins and outs of interactive analysis tools, not mention all computational tools. In the practical use, there are many metrics deserving considerations, which are too many to be discussed here and we will further explore them later, for example, price, number of enhanced documents, after-sales technical support, stability and robustness, operational speed, data volume, concurrent computation ability, formats of input and output files, interface for secondary development, platform portability, and multi-party collaboration ability.

SQL

More accurately speaking, SQL or Structured Query Language is not a tool. SQL is originally introduced by E.J.Codd and implemented by IBM, and nowadays an ANSI and ISO standard computer language. It is the most widely used structured data query and analysis language supported by most vendors, such as Oracle, DB2, SQL Server, Sybase, Informix, MySQL, VF, Access, Firebird, and sqlite. These vendors offer the vendor-characterized syntax that is almost incompatible with each other.

SQL has the largest number of users and the most extensive platform support. The syntax of SQL is close to that of the natural language, making it easy for programmers to learn. But, on the other hand, the drawbacks of SQL are also quite obvious, for example, it is easy to learn a little but hard to grasp it and implement the complicated data analysis. The nested queries of SQL can be used to implement the liner stepwise computation, but it is hard to implement the stepwise computation that is tree-like and easy to decompose and reuse, just like R language and other tools do. In addition, SQL lacks the quantification-oriented functions that are available for Matlab and other tools. To perform the complex computation, SQL users need to resort to the stored procedure developed by programmers with strong technical background. Since SQL also lacks the object-based access mode that is adopted by esProc and other tools, it is also relatively complex for SQL users to represent the multi-table join, and unfavorable for business analyst to study the problem from the business perceptive.

UI friendliness: ★☆☆☆☆

Technical competence requirement: ★☆☆☆☆

Stepwise computation: ★☆☆☆☆

Support for structured data: ★★★★★

Fixed algorithm: ★☆☆☆☆

Target user: Programmer, database administer

Excel

Excel is a business application developed by Microsoft. It is a spreadsheet with intuitive interface, outstanding computation capability and excellent chart tools. In addition, having incorporated VBA, Excel becomes flexible even more greatly. Almost everyone can use it to well meet the need of any industries, any enterprise, or any workspace environment.

Excel is characterized by its high user friendliness. Excel users can name the variable after the cell, just as natural as that, not requiring any effort to define the name like SAS and other tools do. Excel cells are aligned by nature, saving the effort of typesetting. Excel allows for invoking other cells intuitively and calculating automatically, capable to implement the stepwise computation easily. However, on the other side of the coin, the great universality of Excel makes Excel suffers from the relatively poor specialty. With a cell as a unit, the support of Excel for structured data is quite poor. The functionalities of function are rather simple and the representation ability of Excel syntax is not great enough to handle the complex data analysis and specialized scientific computation.

UI friendliness: ★★★★★

Technical competence requirement: ★★★★★

Stepwise computation: ★★★★☆

Support for structured data: ★☆☆☆☆

Fixed algorithm: ★★☆☆☆

Target user: Financial personnel and even those without technical background

Official website: http://office.microsoft.com/en-us/excel/

R

R was originally created by Ross Ihaka and Robert Gentleman at the University of Auckland. It is an open source developing language and software environment running on multiple platforms. R is of object-oriented programming style with the flexible chart plotting ability, inbuilt with several statistics and mathematics analysis capabilities. The greatest application field of R is the bio-information studies, and it is also applied in the econometrics, finance and economics, and humane studies fields.

The most important characteristic of R is that R is free. R is an open source project maintained by the statistician and the mathematician. In addition, it boasts the elegant and agile mechanics of syntax and the open interface for secondary development, so there are a great number of third party packages. But R lacks of the excellent UI interface. Strong technical background and well-grasped expertise are required to use R well.

UI friendliness (the more ★, the friendlier):
★☆☆☆☆

Technical competence requirement (the more ★, the lower level):   ★★☆☆☆

Stepwise computation (the more ★, the more powerful):
★★★★☆

Support for structured data (the more ★, the better support):
★★★☆☆

Fixed algorithm (The more ★, the greater number and more powerful feature):★★★★☆

Target user: statistician, mathematician, scientist

Official website: http://www.r-project.org/

Matlab

Matlab is a business application developed by [MathWorks,](#) Inc. It is an interactive computing environment and fourth generation programming language for numerical computation, algorithm development, and data analysis. You can also use it to plot graphics and charts and create the user interface. Matlab is widely used in the industrial automation design and analysis, and other fields like the image processing, signal processing, communications, and finance modeling and analysis.

Similar to R, Matlab also has a good scalability. It provides the Toolbox style for users to use, review, edit, and share its extended capabilities. Although the third party functions of Matlab are not as many as that of R, Matlab can provide a more superior quality management and some even more powerful functionalities. Plus, regarding its graphic operation interface, Matlab also has an advantage over R .

UI friendliness: ★★☆☆☆

Technical competence requirement: ★★☆☆☆

Stepwise computation: ★★★★☆

Support for structured data: ★★★☆☆

Fixed algorithm: ★★★★☆

Target user: Industrial engineer and statistician

Official website: [http://www.mathworks.com/products/matlab/](http://www.mathworks.com/products/matlab)

esProc

esProc is business desktop application developed by RAQSOFT Inc, specialized in the interactive analysis on the structured data. esProc advocates the free data analysis, requiring relatively low degree of technical competence. It is also renowned for its agile and easy-to-use syntax system. Therefore, esProc is widely adopted by the organizations with a relatively less strong technical background, including most business users, and some users from industrial and financial sectors.

The most distinctive characteristic of esProc is the Excel-style interface for analysis. Which means esProc is of great usability and ideal to achieve the computation goal that is either complex or decomposition-required to compute in steps. With abundant functions for structured data, esProc improves and over-performs SQL in many respects. However, esProc lacks of the fixed algorithm and functions specific to some industries, such as correlation analysis or regression analysis.

UI friendliness: ★★★★☆

Technical competence requirement: ★★★★☆

Stepwise computation: ★★★★★

Support for structured data: ★★★★★

Fixed algorithm: ★☆☆☆☆

Target user: business data analyst, nonprofessional statistician, and finance analyst

Official website: http://www.raqsoft.com/products/esproc

SAS

SAS is a business application developed by SAS Institute Inc. It is a large information system for decision-making support. On one hand, SAS system is quite complex and strict with data, which makes it hard for users to learn, and on the other hand, it is highly precise and trustworthy. SAS is mainly applied in the natural science, economic decision, and enterprise decision. SAS products are also widely applied and adopted in various areas.

Huge, comprehensive, and profound, SAS is characterized with the greatest number of library functions and features as well as the most powerful capabilities in chart plotting for in-depth applications in almost every sector. Although the user interface of SAS has been improved a lot, it is still less friendly than other analysis software.

UI friendliness: ★★☆☆☆

Technical competence requirement: ★☆☆☆☆

Stepwise computation: ★★★★☆

Structured data support: ★★★★☆

Fixed algorithm: ★★★★★

Target user: Statistician, financial specialist, government, and think tank of multinationals

Official website: http://www.sas.com

SPSS

SPSS or PASW is a business application developed by IBM, mainly applied in the statistical analysis, data mining, and decision support. SPSS not only has a user-friendly interface for analysis, but also a set of common and mature statistical procedures, which can satisfy the work requirement of most nonprofessional statistician to the full. SPSS is mainly applied in the communications, medical, finance, and some social science fields. In fact, SPSS is one of the most widely-used professional analysis tools.

SPSS is committed to building the easy-to-implement statistical procedures. Therefore, SPSS has a powerful graphic user interface that is easy for beginners to accept. Compared with syntaxes of other programming languages, the syntax of SPSS is fairly poor and is incapable for the free analysis other than the fixed algorithms. The menu-style interface is also a major obstacle of stepwise computation.

UI friendliness: ★★★★☆

Technical competence requirement: ★★★★☆

Stepwise computation: ★★☆☆☆

Support for structured data: ★★★☆☆

Fixed algorithm: ★★★☆☆

Target user: Business data analyst, nonprofessional statistician, and financial analysis specialist

Official website: http://www-01.ibm.com/software/analytics/spss/

Findings: R is usually used in the normal scientific calculation and highly capable to handle the interactive analysis. But the user interface of R is rather simple and the technical requirement is high. Matlab is similar to R and mostly used in the industrial design. Some algorithms of Matlab are better and more reliable than that of R. esProc is used in the typical business computation and powerful enough to handle the interactive analysis through a friendly interface, only requiring low level of technical requirements on users. But esProc lacks the fixed algorithms for scientific areas. SAS is similar to R and mostly used in the economic decision and natural sciences with a powerful enough ability in interactive analysis. The interface of SAS is rather simple, which results in a very high requirement on user's technical background. SPSS is mostly used in the social science studies, simple to use and powerful, but not flexible enough to process the complex interactive analysis. As the base of business application, SQL can barely handle the interactive analysis but it is quite difficult. Excel is the commonplace for everyone to handle their daily office work, capable to handle the simple interactive analysis job but not the professional problem

Let me conclude this essay by reiterating this point that there is no perfect tool but the tool suites you best, and sometimes you will need to grasp several tools to use in combination.

Reply to this

Threaded Messages (0)