SURVO 76 EDITOR, A NEW TOOL FOR INTERACTIVE STATISTICAL COMPUTING, TEXT AND DATA MANAGEMENT pdf

1: Mittaaminen ja survey-metodiikka (Matematiikan ja tilastotieteen laitos)

Our focus is on the way of working: These questions are relevant in any area of research. Leaving useful traces while working may save a considerable amount of time, and provide better possibilities for other researchers to comprehend the points of a study. These principles are demonstrated with examples using Survo software and its matrix interpreter. The results of a study are not enough. We should also consider and answer the following questions: How did we achieve the results? How did we justify the analyses? How did we correct the errors? Why did we take a particular step? How do we repeat the steps? Errors in the data or in the working process can not be avoided, since they will occur in each research project. Therefore it should be a routine to correct any error when detected, document the correction, make necessary updates, and continue working without disturbing the thinking. A good test for any researcher and any software environment would be to? In general, we should pay more attention to the quality of the working process: How does the work proceed? How well does it get documented? Since our interest is mainly in the dynamic documentation, we would like to ask: What sort of traces are left behind while working? Can we retrace our steps e? The traces represent saved ideas and documented thoughts that are processed while working. Leaving useful traces may save a considerable amount of time. It supports the researcher in backtracking immediately when certain steps of the process have to be repeated. Useful traces help to avoid re-inventing the wheel, and to hold the process together, especially when managing multiple projects simultaneously. They may also provide better possibilities for other researchers to comprehend the points of a study. Gentleman and Temple Lang [7] have introduced the term compendium to refer to documents that are self-contained mixtures of code, text, and data. Publishing compendiums instead of traditional scienti? A proper documentation of the working process published together with accompanying software tools may encourage reproducible research, meaning that the reader can directly reproduce the results employing the methods that are presented in [7]. Regardless of whether the results are published as a compendium or as a traditional scienti? With a good documentation it is easier to get back on the track. The working process depends heavily on the software environment, which may consist of several di? The support from the software environment is needed in the working process including the documentation. This applies for large scale tasks, but also for small details, which tend to appear repeatedly during the working process. In general, we should have? It is preferable to have possibilities for writing free notes and comments nearby the actual operations that are committed. In this sense, the menu-driven environments usually do not provide enough means for dynamic documentation. There should be detailed mechanisms to access the data in various situations. Interactive display of the data is useful for browsing and searching, but the errors should be corrected in a way that also documents the corrections. Otherwise the corrections could be forgotten, when certain steps of the process have to be repeated. A typical situation might be that the original data are modi? Manual corrections could be di? They would slow down the working process and probably also introduce new errors. Automatic documentation means that the software environment leaves useful Leaving useful traces when working with matrices traces without a continuous need to document the working process manually. For example, it is necessary to have good means of naming di? However, cooperation between the environment and the user is needed, since it is also required that the user actually gives describing names. If the environment provides the means, and the user takes advantage of them, it is possible to achieve automatic documentation in subsequent operations. Another way to support automatic documentation is that the operations of the environment are self-documenting as such. Writing too many comments or notes to explain the steps of the working process is not reasonable. It could even be worse than no comments, especially if the code or data are modi? The approach of Survo di? Despite of its long history

SURVO 76 EDITOR, A NEW TOOL FOR INTERACTIVE STATISTICAL COMPUTING, TEXT AND DATA MANAGEMENT pdf

beginning from the s, Survoâ€"at least its recent developmentâ€"is not widely known. Survo supports the documentation of the working process by its editorial interface [11, 12, 14], which replaced the menu-based interface of SURVO 76 in Editorial interface is a distinguishing construct, in which a special text editor takes care of both input and output of operations related to statistical analysis, matrix and other computations, according to the needs of the user. The operations, expressions, and links activated within the text, possibly producing output back in the editor, can be commented inherently in free format. The outputs in turn can be edited, formatted, highlighted, commented, printed, and used as input for other operations. The editorial interface of Survo also enables fruitful cooperation with other programs, such as R [23]. There is even a specialized, user-friendly frontend, which combines the best properties of R and Survo. For example, the statistical and other graphs are Kimmo Vehkalahti by default plotted as meta? The rotations in factor analysis can be performed in an interactive graphics window as a sequence of two-dimensional rotations. The matrix operations of Survo are carried out by a matrix interpreter, which is closely integrated in the statistical, graphical, and other operations of the system through the editorial interface and the data and matrix structures [18, pp. Features of the matrix interpreter, such as automatically inherited row and column labels of the matrix objects, support the documentation of the working process. Mustonen [20] has demonstrated matrix computations in Survo, beginning from elementary operations and extending to advanced applications and programming. The data sets and sucros or Survo macros [16] utilized in the paper are included in the distribution package of Survo. Obviously, this idea is equivalent to the principles that have been recently termed reproducible research [7], since the accompanying data sets and sucros allow the user to directly or even automatically reproduce the results and employ the methods that are presented in the paper. The user may also freely modify the premises and make alternative analyses and simulations using the ready-made working schemes provided in the saved Survo jobs. Reproducibility does not depend on whether we are working with matrices or not, as the common factor of all activities in Survo is the editorial interface. Working in Survo typically results in saved jobs or documents that not only give detailed descriptions of the working processes but also facilitate repeating the processes more or less automatically. Hence, it can be said that the idea of reproducible research has been a regular way of working in Survo since the invention of the editorial interface in [11]. The core structures of Survo have not been modi? Its programming libraries, including a set of matrix functions have been described in [17] with the details of the data and matrix structures. The libraries are updated synchronously along the system development, and they are freely available. Survo is a modular system which can be extended without limits [17]. The interaction between the community of Survo users and the developers has been rewarding both for the users and the developers [18, p. The developers appreciate the feedback and suggestions, while the users appreciate the continuous development process. However, because of the continuous development process, the concept of version is not very signi? According to one of the major development principles of Survo, new functions are added to the system in a way that the existing ones are not changed, since that could potentially cause harm for any user-de? By adding new options, the functions can still be extended without limits and without disturbing existing usage. Applying the above principle has guaranteed a full backward compatibility between the versions of Survo based on the editorial interface [21]. Therefore, the user Leaving useful traces when working with matrices never needs to re-learn how to use Survo. The interface is the same, although the technical environment, such as the operating system has changed and might change again. In our terms, Survo is an environment where the traces left behind can be retraced also in the future. A traditional research paper can not give a true picture of any dynamic contents, but hopefully the following examples will provide an idea of the documentation possibilities in the Survo environment. Before turning to a particular example, we brie? Factor analysis was originally invented in by Spearman [25], and developed further both by psychologists and statisticians. The mistrust in factor analysis has persistently been present in the statistical literature. In their introductory book of multivariate analysis, Chat? However, this mistrust is mainly caused by misconceptions. But, as Mustonen [19, p. The problem is that the conclusion is based on false assumptions: Francis had believed that Kimmo Vehkalahti the given

factor pattern would represent a simple structure [28] which should be reproduced by factor analysis. However, the structure is not simple in this sense, which makes it quite impossible to reproduce by standard methods [19, p. The meta information and documentation around the numerical elements consist of 1 general comments, including the internal name of the matrix, 2 row labels, and 3 column labels. The structure would also allow rowwise comments, which are typically used in vectors giving some scalar results e. All information is included in the matrix object when it is saved as a matrix? MAT in the current working directory. In subsequent operations, the matrix is referred to by using this external name. The internal name of the matrix will be updated automatically by each matrix operation. Comments may be written all around the commands, even on the same line when separated with an isolated slash. They are ignored by the interpreter the MAT commands , but they might be valuable for the user in retracing the steps of the work in the future.

SURVO 76 EDITOR, A NEW TOOL FOR INTERACTIVE STATISTICAL COMPUTING, TEXT AND DATA MANAGEMENT pdf

2: Publications related to Survo

SURVO 76 EDITOR, a new tool for interactive statistical computing, text and data management (release 2). Research Report No. Research Report No. Department of Statistics, University of Helsinki.

Nuorempana otin joitain paperikuvia pokkareilla. Survo on kuin erikoisvarusteltu askarteluhuone. Osoitteesta kirjat saa noin e yhteishintaan. Pisteen koko on verrannollinen kuvien keskikokoon. Seuraavassa askartelen Survon keinoin tuon ajanjakson kimpussa: Helppoa, kunhan sopiva piirtodata on ensin muodostettu. Otsikkoon innoittanut kansikuva ei tosin ole Japanista vaan Espanjan aurinkorannikolta. Asiakaspalvelussa Japani on huippumaa. Ilmojen suhteen oli onnea: Suosittelen todellakin tutustumista Japaniin, eritoten Tokioon. Se on rakennettu samaan tapaan kuin sivun 3 kuva. DV2 on tavallinen tekstitiedosto. Ne on koottu LaTeX. Survon tekstieditorissa koodit saadaan niin piiloon kuin halutaan. Lienee makuasia, miten laajasti haluaa koodaukset piilottaa. Periaatteet olivat samat, vaikka kyse olikin HTML: Lopputulos olisi siis ollut sama: Haasteista narinaa Tulokset ratkaisevat, sanotaan. Vaiheet on voitava toistaa samantien, tarvittaessa automaattisesti. Addison-Wesley, Boston, toinen painos. A Document Preparation System. User s Guide and Reference Manual. Frank Mittelbach ja Michel Goossens Factor Analysis and the Reliability of Measurement Scales. Effects of measurement errors in predictor selection of linear regression model.

SURVO 76 EDITOR, A NEW TOOL FOR INTERACTIVE STATISTICAL COMPUTING, TEXT AND DATA MANAGEMENT pdf

3: Microsoft Excel - Wikipedia

Editorial mode is a new form of interactive computing enabling the user to perform mathematical, statistical and graphical operations in connection with normal text editing. Various aspects of the editorial approach are considered through examples. Mustonen, S., SURVO 76, a statistical data.

Release of Micro Data. How to do this ideally? Methods for achieving equivalence of samples in cross-national surveys. Muennich, Ralf; Schuerle, Josef; Dacseis Project Research Papers under Workpackage 3. Quatember, Andreas; Haslinger, Alois; Analysis of National Surveys. Dacseis Project Research Papers under Workpackage 2. Dacseis Project Research Papers under Workpackage 4. Integrated Modelling Approach to Imputation. Working papers of the Office for National Statistics. Evaluation of Imputation Software. Juulia; Partanen, Jukka M. Television viewing is associated with psychiatric symptoms in -year-old children. Matrices as building blocks of measurement framework. Enhancing the documentation by leaving useful traces. Factor analysis and the reliability of measurement scales. Statistical Simulations in the Web. Linear Regression Model with Measurement Framework. Leaving useful traces when working with matrices. Two-Phase Sampling and Weighting. Learning by Co-operating and Doing. Experiences on developing business register systems and exploiting business data. Amrads Final Conference, November. The junction between external data and statistics data. Is it possible to optimise? Methodological Approach to Harmonisation of Multinational Surveys. Exploitation of Statistical Business Data. Overview to Create, validate, maintain and utilise complex data. Published in Pre-Proceedings, pp. Proceedings of the American Statistical Association. Weighting and Auxiliary Variables in Sample Surveys. Methodes, modeles, applications, nouvelles approaches". Matrix computations in Survo. Longitudinal study using synthetic units of Finnish manufacturing. Nonresponse in Survey Research, pp. Approches nouvelles en sondage". Research Reports , Ponderation et variables auxiliaires dans les enquetes par sondages. Colloque Francophone sur les Sondages, Survo as an environment for statistical research and teaching. Statistical Properties of Reliability Measures. University of Tartu, Estonia, pp. International Perspectives on Nonresponse. Statistical Reports of the Nordic Countries 66, Techniques and uses of enterprise panels. Luxembourg, November International Comparability of Business Surveys: Some Key Factors based on European Experiences. Book of invited papers, Statistical Methodology and the Principle of Subsidiarity: Total quality and computer-assisted interviewing: Essays on Blaise New trends in probability and statistics. Multivariate statistics and matrices in statistics. European Enterprise Panels Network: Luxembourg, February International Association for Statistical Education: Proceedings of the first scientific meeting, University of Perugia, August , Editorial interface in Statistical Computing and Related Areas. Teaching statistics in a science centre. General Environment for for Statistical Computing. Gustav Fischer Verlag, Stuttgart. Editorial approach in statistical computing. Statistical computing based on text editing. Proceedings in Computational Statistics, ed. Statistical computing with a text editor. Walter de Gruyter, Berlin. Proceedings in computational statistics, ed. Fitting alternative regression models to heterogeneous data. A theoretical overview for variance estimation in sampling theory with some new techniques for complex estimators. Doctoral dissertation, Department of Statistics, University of Helsinki. On Reliability of Composite Scales: An essay on the structure of measurement and the properties of the coefficients of reliability - a unified approach. Oppimiseen kannustava ilmapiiri sulautuvassa opetuksessa. Proceedings of the 14th International Workshop on Matrices and Statistics. Massey University, Auckland, New Zealand, Results of Recent Econometric Research. Laaksonen, Seppo ed. The Evolution of Firms and Industries. Statistics Finland Research Reports International Perspectives in Nonresponse. Kirjavainen et al Housework Time in Bulgaria and Finland. Modelling for Housework Time Cross Nationally. Housework Time In Bulgaria and Finland. Wage from Work and Gender. Statistics Finland, Studies , Opetusmateriaalia Vehkalahti, Kimmo Data-analyysi I - Survo-kurssi. Tilastotieteen laitos, Helsingin yliopisto. Department of Statistics, University of Helsinki. Survo-opas Data-analyysi I -kurssia

SURVO 76 EDITOR, A NEW TOOL FOR INTERACTIVE STATISTICAL COMPUTING, TEXT AND DATA MANAGEMENT pdf

varten pp.

SURVO 76 editor, a new tool for interactive statistical computing, text and data management. Research report / University of Helsinki. Research report / University of Helsinki. Department of Statistics no.

Statistical computing requires appropriate tools. We describe the integrated editorial environment for data analysis and related tasks developed since the s under the name of Survo and its recent open source R package implementation known as Muste. As examples of the editorial approach in Muste, we consider a statistical application on influence curves for a correlation coefficient as well as on simulation and contour curves of bivariate Gaussian data. Muste, Survo, R project, statistical software, text editor, user interface, data analysis, bivariate normal distribution, history of statistical computing 1 Introduction Statistical computing is not possible without appropriate tools. In practice, any statistician working with concrete data requires software packages that help to analyse data with computers. Quite recently, the R project [21] has become an increasingly popular tool that allows enough flexibility even for very specialized tasks requiring novel approaches to data analysis [4]. R is an open source software package that has roots in a research project of Robert Gentleman and Ross Ihaka in the s [6]. R originated as a Scheme variant, but has developed into a dialect of S language that was introduced in the Bell Labs by John Chambers, Rick Becker and others in [2]. Another example of a flexible tool is Survo, an integrated environment for statistical com- puting and related areas, developed since the s by Seppo Mustonen [1, 9, 13, 15, 17, 20]. The central feature of Survo is the editorial interface that was introduced in Since then, Survo has allowed statistical computing and data analysis to be an inseparable part of writing and reporting the results in a research process. Somewhat similar ideas have been described more specifically under different headings such as literate programming [7], re- producible research and compendium [5]. Certain properties of using statistical programs from a Muste editorial computing text editor are also present in the Emacs Speaks Statistics approach [22], TeXmacs platform [24] and Mathematica notebooks [25]. However, the main distinction is that Survo has always provided a general, integrated and self-documenting environment for statistical computing and related areas in one stand-alone software program. The aim of this paper is to describe how Survo and R have been combined in the Muste project that was initiated in by Reijo Sund. So far we have successfully implemented the key parts of the Survo system as an open source R package Muste. We begin with a brief review of the long history of Survo and the short history of Muste. As examples of the editorial approach in Muste, we consider a few small statistical applications. There, during a coffee break on a hot summer day in the late August, Mustonen and his colleague Martti Tienari told professor Olli Lokki about the ideas concerning the future of data processing developed in the Finnish Cable Works. Professor Lokki was very enthusiastic about the more or less specific dreams about how the things related to statistical computing should be arranged. During those discussions the aim to develop a whole statistical programming language was presented for the first time. Mustonen extended these ideas with Tienari and Timo Alanko during the years â€" and prepared a first proposal for a statistical programming system SURVO The system allowed the user to define the analyses to be performed using a simple command language [1]. Editorial approach and general computing environment Since the very first ideas of Survo in the s, the challenges of integrating the various activities and the interactivity between the user and the computer have had an important role in the development of Survo. While the computers of the s were only able to satisfy the wishes to a slight degree, the computers of the next decade offered much better options. It provided an interactive menu-driven environment in addition to a command language, and its use was like a conversation with the computer [9, 10]. This approach means that the interactivity in data analysis is achieved by working within a text editor that allows a free mixture of natural language, data, activatable commands and results [11]. Writing of music sheets certainly was not the main purpose of an interactive statistical program, but these experiments convinced Mustonen to develop the idea further. Since then, the idea and features of the editorial approach have been continuously extended in

many ways. Muste is an R implementation of Survo The main motivation for the Muste project was to transform the unique or otherwise useful properties of Survo system into the form of an open source multiplatform software. As the R project had become very popular, offering an appealing environment for implementing open source statistical software for multiple platforms, it was a natural choice to test the possibilities of producing a Survo-look-alike functionality in that context. It quickly turned out that the text widgets available via the tcltk package [3], a base part of R, would allow the implementation of the Survo editor in Muste. The initial tests in were promising and Seppo Mustonen kindly provided all the necessary parts of the Survo source code written in C for the Muste project. Due to this and the fact that virtually all platform specific parts of Survo had been isolated and thus made easily replaceable, the only functions that had to be re-programmed from scratch, were the input and output functions. Most of the functionality of the Survo system has already been included in Muste. Muste is available for anyone interested from the R-Forge development platform: When the editor is launched with muste , a new editor window, similar to the one presented in Figure 1 appears. Muste editor works as a text editor in which the user can freely write text on an edit field. The text may contain commands, operations, specifications and other parts that Muste will interpret as the user activates them, by the Esc key or the mouse double-click. The header line displays the current date and time, the working directory, the size of the edit field, and certain status information. Technically, the shadows are another layer of text reflected in different colors in the edit field. In general, the shadows may have both visual and structural purposes in various operations. Usually, there is no need to highlight this distinction explicitly. Instead, the user may choose whatever s he wants to highlight with the shadows. The editorial approach gives more versatile possibilities for mixing text, commands, data and results, because any elements can be freely combined in the edit field, with their exact meaning to be interpreted dynamically, depending on the context at the moment of the activation. In Figure 1, both data and commands can be seen, in addition to free-form text. These data will then be used to produce the results that immediately appear on lines 16€" Editorial arithmetics and conversions Muste editor contains a powerful tool for making many types of calculations with a flexible interface. The example that continues in Figure 2, shows how to access the results of the CORR command in editorial arithmetics. The results are written to the edit field by activating the calculation after the equal sign. In addition to a large selection of various conversion options available in Muste, the user can easily add new ones by defining a few simple rules. Certain special transformations, such as numbers with different bases, roman numerals, written words or factors of an integer are included by default. As a slightly more complicated example, we consider adding one new observation to the data and study its effect on basic statistics. In order to use the equation 2 in editorial arithmetics of Muste, it is defined on line 37 of the example in Figure 2. On line 42, we define a function that looks for a value for the weight fulfilling the conditions, using a grid search over integer values on a reasonable interval. Influence curves for the correlation coefficient Our example continues in Figure 3, where we proceed with the correlation coefficient of two variables, x and y. They are then applied in editorial arithmetics on lines 56€" In addition, they are used below the calculations to plot a graph. Weight 90 80 70 60 50 40 Height Figure 4: Thus, the final graph depicts contours of r with increments of 0. Contour ellipses for bivariate Gaussian data We conclude the example by simulating data from a bivariate Gaussian distribution and by drawing contour ellipses for the simulated data. Muste includes a specialized command MNSIMUL for simulating data from multinormal distribution, but here we construct the variables manually, using formulas for bivariate Gaussian distribution presented by Seppo Mustonen [16, 19], who derived them as generalizations to the well known Box€"Muller formulas. Let U1 and U2 be independent variables from the uniform distribution on 0, 1. The random variables are generated by the VAR command that utilizes the definitions of the Mustonen formulas on lines 73€" The means, standard deviations and the correlation coefficient come from the DECA data introduced in the beginning of the example, in Figure 1. Here we, however, plot the contour ellipses manually. The actual PLOT command for the contour ellipses is on line The EPS command on line 84 combines the plots of the data and the contour curves as one PostScript file. The final graph is displayed in Figure 6. In order to use

SURVO 76 EDITOR, A NEW TOOL FOR INTERACTIVE STATISTICAL COMPUTING, TEXT AND DATA MANAGEMENT pdf

different parameters with the equations, we have drawn a border line on line It restricts the scope of all specifications above and below it. However, on line 82 we have locally extended the default scope with the SPECS specification, allowing all the lines from 73 to 89 marked with symbols a and b in the control column to be included when the Muste editor scans the specifications. Technical details like this demonstrate how the design of a Muste work scheme can resemble, at least to some extent, formal programming. Currently, most parts of Survo are implemented in Muste. In this paper, only a few rather simple examples of the editorial approach in Muste were briefly described. In this respect, Muste is a software within a software, although it is as integral a part of R as any other R package. Muste can also be used as a flexible script editor or even more generally, a GUI for R. In addition, Muste provides advanced tools such as Survo data file support and matrix interpreter. With these tools several types of data preprocessing tasks may be easier to perform with Muste than with the traditional R. As Muste data sets are easily transformed to and from R data frames, the data analysis can be performed by alternating between the Muste functions and any R packages providing the methods needed for particular tasks. Compared with the usual interfaces, such as the default interface of R, a great advantage of using Muste is the self- documenting feature of its editorial approach, inherited from the original Survo editor from the beginning of s. A detailed exposition of the above-mentioned ways of working is excluded from this paper and will be described elsewhere. Kimmo Vehkalahti has participated in the development of Survo and Muste as well as in the critical revision of the manuscript. Seppo Mustonen is the developer of Survo and his contributions to the Muste project have been extremely significant as most of the functionality in Muste is directly based on his original source codes of Survo. He has participated in the critical revision of the design and contents of the manuscript. We would also like to thank two anonymous referees for useful comments. Bibliography [1] Alanko, T. Springer Series in Statistics and Computing. R News, 1, 27â€" The R Journal, 1, 5â€" Journal of Computational and Graphical Statistics, 16, 1â€" A Language for Data Analysis and Graphics. Journal of Computational and Graphical Statistics, 5, â€" The Computer Journal, 27, 97â€" National Elliott computer application program LM Elliott Computing Division, Bore- hamwood. Department of Statistics, University of Helsinki. Scandinavian Journal of Statistics, 8, â€" Univer- sity of Helsinki, Department of Statistics.

I d. I P.,.1ì}" 4" c Jr 6 ffifr3$* T'ffiF* Tffi trntrmductÃ¥nn Ã¥n inÃ¥rmdurtary exsmplp Tha us# #S trdÃ¥t trsy* Tex t sd i"*Ã¥n g 4,.1" C Ssar Ã¥n g the ffd iÃ¥ f *e 3"d 4"Ã¤".*

To save the chart and spreadsheet save as. XLC is not supported in Excel or in any newer versions of Excel. Dialog Used in older versions of Excel. Template A pre-formatted spreadsheet created by the user or by Microsoft Excel. The most common are Dynamic Data Exchange: It is very common in financial markets, being used to connect to important financial data services such as Bloomberg and Reuters. This may take on the form of "embedding" where an application uses another to handle a task that it is more suited to, for example a PowerPoint presentation may be embedded in an Excel spreadsheet or vice versa. Excel can accept data in real time through several programming interfaces, which allow it to communicate with many data sources such as Bloomberg and Reuters through addins such as Power Plus Pro. Although it is easy for users to create such links, programming such links reliably is so difficult that Microsoft, the creators of the system, officially refer to it as "the protocol from hell". Network DDE Extended the protocol to allow spreadsheets on different computers to exchange data. Starting with Windows Vista, Microsoft no longer supports the facility. RTD although in many ways technically superior to DDE, has been slow to gain acceptance, since it requires non-trivial programming skills, and when first released was neither adequately documented nor supported by the major data vendors. These include opening Excel documents on the web using either ActiveX controls, or plugins like the Adobe Flash Player. ExcelPackage is another open-source project that provides server-side generation of Microsoft Excel spreadsheets. Excel Services is a current. Excel spreadsheets can be accessed from Python with xlrd and openpyxl. Password protection Microsoft Excel protection offers several types of passwords: Password to open a document [43] Password to modify a document [44] Password to unprotect worksheet Password to protect workbook Password to protect the sharing workbook [45] All passwords except password to open a document can be removed instantly regardless of Microsoft Excel version used to create the document. These types of passwords are used primarily for shared work on a document. The only type of password that can prevent a trespasser from gaining access to a document is password to open a document. The cryptographic strength of this kind of protection depends strongly on the Microsoft Excel version that was used to create the document. In Microsoft Excel 95 and earlier versions, password to open is converted to a bit key that can be instantly cracked. As regards services which use rainbow tables e. Password-Find , it takes up to several seconds to remove protection. In addition, password-cracking programs can brute-force attack passwords at a rate of hundreds of thousands of passwords a second, which not only lets them decrypt a document, but also find the original password. The situation changed fundamentally in Excel , where the modern AES algorithm with a key of bits started being used for decryption, and a 50,fold use of the hash function SHA1 reduced the speed of brute-force attacks down to hundreds of passwords per second. In Excel , the strength of the protection by the default was increased two times due to the use of a ,fold SHA1 to convert a password to a key. Microsoft Excel Viewer Microsoft Excel Viewer is a freeware program for viewing and printing spreadsheet documents created by Excel. There is not a current version for the Mac. Microsoft has announced some of these issues are addressed in Excel In the case of excessively large results, Excel will return the error warning NUM! Date range Excel supports dates with years in the range , except that December 31, can be entered as 0 and is displayed as 0-jan Converting a fraction of a day into hours, minutes and days by treating it as a moment on the day January 1, , does not work for a negative fraction. A similar problem occurs when a text happens to be in the form of a floating point notation of a number. In these cases the original exact text cannot be recovered from the result. In the case of entering gene names this is a well known problem in the analysis of DNA , for example in bioinformatics. The problem was first described in You cannot open two documents with the same name, even if the documents are in different folders. To open the second document, either close the document that is currently open, or rename one of the documents. Numeric

precision in Microsoft Excel Excel maintains 15 figures in its numbers, but they are not always accurate: Despite the use of figure precision, Excel can display many more figures up to thirty upon user request. But the displayed figures are not those actually used in its computations, and so, for example, the difference of two numbers may differ from the difference of their displayed values. Although such departures are usually beyond the 15th decimal, exceptions do occur, especially for very large or very small numbers. Serious errors can occur if decisions are made based upon automated comparisons of numbers for example, using the Excel If function , as equality of two numbers can be unpredictable. Although this number has a decimal representation that is an infinite string of ones, Excel displays only the leading 15 figures. In the second line, the number one is added to the fraction, and again Excel displays only 15 figures. In the third line, one is subtracted from the sum using Excel. This is because Excel calculates with about half a digit more than it displays. Excel works with a modified version of the IEEE specification. See the main article for details. Besides accuracy in user computations, the question of accuracy in Excel-provided functions may be raised. Particularly in the arena of statistical functions, Excel has been criticized for sacrificing accuracy for speed of calculation. IsText " in VBA , incorrectly returns "false". Versions Early history Microsoft originally marketed a spreadsheet program called Multiplan in Microsoft released the first version of Excel for the Macintosh on September 30, , and the first Windows version was 2. This accomplishment solidified Microsoft as a valid competitor and showed its future of developing GUI software. Microsoft maintained its advantage with regular new releases, every two years or so. Microsoft Windows Excel 2. Versions prior to 2. This included a run-time version of Windows. The magazine stated that the port of the "extraordinary" Macintosh version "shines", with a user interface as good as or better than the original. VBA is a powerful addition to the application and includes a fully featured integrated development environment IDE. Macro recording can produce VBA code replicating user actions, thus allowing simple automation of regular tasks. The automation functionality provided by VBA made Excel a target for macro viruses. This caused serious problems until antivirus products began to detect these viruses. Microsoft belatedly took steps to prevent the misuse by adding the ability to disable macros completely, to enable macros when opening a workbook or to trust all macros signed using a trusted certificate. Internal rewrite to bits. Almost no external changes, but faster and more stable. This was a major upgrade that introduced the paper clip office assistant and featured standard VBA used instead of internal Excel Basic. It introduced the now-removed Natural Language labels. This version of Excel includes a flight simulator as an Easter Egg. This was a minor upgrade, but introduced an upgrade to the clipboard where it can hold multiple objects at once. The Office Assistant, whose frequent unsolicited appearance in Excel 97 had annoyed many users, became less intrusive. Minor enhancements, most significant being the new Tables. This release was a major upgrade from the previous version. Similar to other updated Office products, Excel in used the new Ribbon menu system. This was different from what users were used to, and was met with mixed reactions. One study reported fairly good acceptance by users except highly experienced users and users of word processing applications with a classical WIMP interface , but was less convinced in terms of efficiency and organisation. Also added was an improved management of named variables through the Name Manager, and much improved flexibility in formatting graphs, which allow x, y coordinate labeling and lines of arbitrary weight. Several improvements to pivot tables were introduced. To illustrate, the number of rows was now 1,, and columns was 16, ; the far-right column is XFD. Minor enhancements and bit support, [84] including the following: Multi-threading recalculation MTR for commonly used functions Improved pivot tables.

SURVO 76 EDITOR, A NEW TOOL FOR INTERACTIVE STATISTICAL COMPUTING, TEXT AND DATA MANAGEMENT pdf

Survo: An Integrated Environment for Statistical Computing and Related Areas () Survo, Arvo. Survo 76 editor, a new tool for interactive statistical computing, text and data management (Research report / by Seppo Mustonen ().

The list below includes the publications related to various phases of Survo development during the decades. On Distance Distributions in Networks. Alkoholipoliittisen tutkimuslaitoksen tutkimusseloste N: On Statistical Multivariate Methods. Research Reports of the Alcohol Policy Institution. Multiple Discriminant Analysis in Linguistic Problems. Statistical Methods in Linguistics, 4, Monistesarja, Tampereen yliopiston tietokonekeskus, Moniste n: Musiikki [a Finnish journal of Musicology] 3, Department of Statistics, University of Helsinki. Fitting alternative regression models to heterogeneous data. Statistical computing using a mini computer. SURVO 76, a statistical data processing system. SURVO 76, an interactive statistical data processing system for a desk computer. Arkhimedes 30 1 , Arkhimedes [a Finnish journal of Physics and Mathematics] 30 1 , Teoksessa Tilastotieteen sovellutuksista, toim. Vartia, Helsingin yliopiston tilastotieteen laitoksen opetusmonisteita n: On Statistical Computing, in Applications of Statistics, ed. Vartia, University of Helsinki, Department of Statistics, pp. On Interactive Statistical Data Processing. SURVO 76 -uutisia, SURVO 76 program descriptions. Statistical computing with a text editor. Computational Statistics, Festschrift dedicated to Dr. Walter de Gruyter, Berlin, New York. Statistical computing based on text editing. Two programs for time series analysis. SURVO 84 - interactive system for statistical computing, graphics and text processing. Helsingin yliopisto, tilastotieteen laitos. University of Helsinki, Department of Statistics, pp. University of Helsinki, Department of Statistics. Dimensio, 52 5 , Dimensio, a Science Journal. Analysis of variance and covariance. Helsinki School of Economics, Department of Methodology. Tampereen yliopisto, Matemaattisten tieteiden laitos, Moniste B Exercices of Introduction to Statistics II. A singular value decomposition of matrices in a space with an indefinite scalar product. Mathematica Dissertationes, 79 40 pp. General Environment for Statistical Computing. Recent development of Survo. Programming sucros in Survo. Network version of Survo. Editorial interface in Statistical Computing and Related Areas. Survo-ohjelmiston versio 4 julkistettu, Version 4 of Survo released, 14 Sep Version 4 of Survo on A4. Analyzing linguistic data with Survo. Freeware version of Survo: On the recent development of Survo. Examples of the use of Survo in teaching regression concepts. Survo-opas Data-analyysi I -kurssia varten. Survo Guide for Data analysis I. Survo Systems, Helsinki, pp. On the version 5 of Survo. Helsingin yliopisto, psykologian laitos. University of Helsinki, Department of Psychology. A measure for total variability in multivariate normal distribution. New Survo on the way! Survo as an environment for statistical research and teaching. Further examples of the use of Survo in teaching regression analysis. Matematiikan, tilastotieteen ja filosofian laitos, Tampereen yliopisto, B Matrices for a Statistician. Matrix computations in Survo. Also available in PDF format , originally distributed on paper for the workshop participants. Windows version of Survo on the way! Pro gradu -tutkielma, Helsingin yliopisto, Tilastotieteen laitos. Statistical methods for modelling of dynamic patient populations. Reliability of measurement scales: The new Windows version of Survo. Data-analyysi I -Survo-kurssi. Data-analysis I - Survo course. Minimum Description Length based model selection in linear regression. Logarithmic mean for several arguments. Influence curves for the correlation coefficient. University of Helsinki, Department of Mathematics and Statistics. Computing Environment for Statistical Research and Teaching. Leaving useful traces when working with matrices. Research Letters in the Information and Mathematical Sciences, 8, On certain cross sum puzzles. Luvut, num3rot ja kuvat. Figures, numb3rs and graphs. Effects of measurement errors in predictor selection of linear regression model. Tilastotieteen pro gradu -tutkielma. Tampereen yliopisto, Informaatiotieteiden tiedekunta, Matematiikan ja tilastotieteen laitos. Methodological perspectives for register-based health system performance assessment: Developing a hip fracture monitoring system in Finland.

7: [PDF] Statistics An Introduction Using R Download eBook for Free

[11] Mustonen, S. (b) SURVO 76 EDITOR, a new tool for interactive statistical computing, text and data management. Research Report No. Department of Statistics, University of Helsinki.

R is supported by various packages to compliment the work done by control structures. R offers wide range of packages for importing data available in any format such as. To import large files of data quickly, it is advisable to install and use data. R has in built plotting commands as well. They are good to create simple graphs. But, becomes complex when it comes to creating advanced graphics. Hence, you should install ggplot2. These packages are dplyr, plyr, tidyr, lubridate, stringr. Check out this complete tutorial on data manipulation packages in R. For modeling, caret package in R is powerful enough to cater to every need for creating machine learning model. However, you can install packages algorithms wise such as randomForest, rpart, gbm etc Note: You might like to check this interesting infographic on complete list of useful R packages. But before you proceed. Then type, library swirl to initiate the package. And, complete this interactive R tutorial. If you have followed this article thoroughly, this assignment should be an easy task for you! In case you find anything difficult to understand, ask me in the comments section below. Data Exploration is a crucial stage of predictive model. This stage forms a concrete foundation for data manipulation the very next stage. In a data set, the response variable y is one on which we make predictions. Refer to image shown below Predictor Variable a. The predictive model is always built on train data set. This data always contains less number of observations than train data set. Right now, you should download the data set. Take a good look at train and test data. Cross check the information shared above and then proceed. To check if the data set has been loaded successfully, look at R environment. The data can be seen there. Test data should always have one column less mentioned above right? This can be done by using: Many data scientists have repeatedly advised beginners to pay close attention to missing value in data exploration stages.

ment for statistical computing and related areas known as Survo has been fully incorporated into R as an open source package muste freely available from the R-Forge repository. Survo has been developed since the early s by professor Seppo Mustonen.

Various forms of the Survo system are essential parts in this tradition. The first Survo systems were created in sixties Alanko, Mustonen, Tienari and seventies Mustonen SURVO 76 was one of the first interactive statistical packages. All this happened before the current microcomputer age. These ideas are based on a concept of editorial approach. In this approach all functions of the system are controlled by a specific text editor which distributes the tasks between various independent program modules. The Survo editor is the center of all activities and the system as a whole is a general environment for various tasks related not only to statistical analysis and computing. In fact, we have extended functions of Survo to many areas which are important in statistical research and teaching of statistics. Fundamentally using Survo is like working with a combined word processor and spreadsheet program with extended capabilities to various directions. Thus in Survo one can maintain the whole statistical research process. For this purpose Survo includes functions for data input and screening, general data management, statistical graphics and analysis, matrix computations, making reports in printable form, desktop publishing, etc. Survo also includes a powerful macro language which enables making of various expert applications by combining automatically and conditionally ready-made functions of the system. The same technique can be used for creating teaching programs on any topic, for example, on certain statistical methods. Survo macros are called sucros. Plenty of teaching programs have been made as sucros on topics related to statistics. Survo provides also means for making hypertext applications. For example, a basic course on multivariate statistical methods has been made by the first author in Finnish. It has been published as a textbook, but all the text, formulas, numerical examples, data sets, and sucros related to the topic are available in electronic form when using Survo. Thus when giving the course both text and examples can be shown on the screen in the classroom. The teacher can easily modify the examples and repeat various stages of analyses and simulations during the lecture. In a short paper it is impossible to render any representative examples about how Survo is used in true research and teaching situations. However, we hope that the following tiny application related to problems in factor analysis FA illustrates something essential. In the textbook on multivariate analysis Seber , pp. Francis was studying various artificial factor structures of ten variables and 2 or 3 factors by creating samples typically of size 50 and trying to detect the original factor pattern by the standard methods, for example, by maximum likelihood factorization and varimax rotation. On the basis of these experiments Francis as quoted by Seber came to very negative results. We think that such claims are exaggerated since many of the experiments of Francis are misleading. To make this evident, we shall re-evaluate one typical experiment Model V of Francis. In the sequel we simply show what the user has typed in the edit field and what Survo has given as results given here in gray shading. The commands activated by the user are displayed as white text on black background. Basically, everything in Survo is carried out in an edit field which corresponds to a spreadsheet but has also capabilities of a word processor. The user types text and commands in this working area. When a command is activated, the editor program passes the task to a suitable program module. The results are automatically written partly in the same edit field in legible form and partly into files numerical results in double precision. Thus the following example should be seen as a pale projection of a dynamic process where the actions of the user and the computer are efficiently interlinked. As an integrated environment Survo is suitable for various research and teaching applications. As an example some problems related to factor analysis are considered. This example shows how a work process is documented by combining statistical computations and text editing.

SURVO 76 EDITOR, A NEW TOOL FOR INTERACTIVE STATISTICAL COMPUTING, TEXT AND DATA MANAGEMENT pdf

9: Chernoffin naamat ja Suomen vointi (hyvin? pahoin?)

SURVO 76 was one of the first interactive statistical packages. All this happened before the current microcomputer age. The present Survo (SURVO 84C) was started about 13 years ago from new ideas developed and tested in a rudimentary form already in SURVO

As notes they are schematic, but quite detailed and very well arranged. Most topics of various phases of there research process are covered: Social research methods A free "teach yourself" tutorial teaching internet skills for those working with social statistics. Ideal for independent learning or for use by teachers and lecturers in their courses. It takes around an hour to do and has quizzes and interactive exercises to lighten the learning experience. This is just one of over 60 tutorials within the Virtual Training Suite. All of the tutorials are written and reviewed by a national team of lecturers and librarians from universities across the UK. Social Research Update SRU A quarterly electronic journal which covers new developments in social research, one per issue. Methods of doing social research are continually advancing. New developments in information technology enable complex analyses to be carried out on a personal computer. Laptop and computer assisted telephone interviewing change fundamentally the way in which surveys are done. New statistical techniques are being developed which are better suited to the kinds of data which social researchers generally deal with. New approaches to qualitative data raise new theoretical and ethical problems. In face of all this change, social researchers, pressurised to produce results, often do not have time to keep up with the latest developments. It is this gap that Social Research Update attempts to fill. The site serves as the gateway to the Electronic Paper Archive. From here you can view abstacts of conference and working papers online and download papers for local printing. It is also the gateway to the newsletter of the section, The Political Methodologist and to the official journal of the section Political Analysis. Sociological Methodology An annual volume on methods of research in the social sciences. Sponsored by the American Sociological Association, its mission is to disseminate material that advances empirical research in sociology and related disciplines. Chapters present original methodological contributions, expository statements on and illustrations of recently developed techniques, and critical discussions of research practice. Sociological Methodology seeks contributions that address the full range of problems confronted by empirical work in the contemporary social sciences, including conceptualization and modeling, research design, data collection, measurement, and data analysis. Qualitative Research Action Science Network Aims to accurately describe and efficiently demonstrate the theory and practice of action science. Action science is a strategy of organizational development which increases the skills and confidence of individuals in groups to create any kind of organization. Action science was defined primarily by Chris Argyris with important help from Donald Schon and others.

SURVO 76 EDITOR, A NEW TOOL FOR INTERACTIVE STATISTICAL COMPUTING, TEXT AND DATA MANAGEMENT pdf

La campanella sheet music violin Precipice (Macmillan UK Audio Books) Maya Sangh And The Valley of the White Ones Indian reservation system Planning a Christian funeral In the service of Congress, 1775-1783 Hoist Your Sails Run Sex for the clueless Usage and usability assessment Modernising hunger The Vice Presidents task force. The warrens player map Episodes from the Early History of Mathematics (New Mathematical Library) Part 7 : The Washington years Frm schweser 2014 Its a matter of philosophy : remedies x 10 Problems of the Indian creative writer in English Electron donating group list Shakespeare twelfth night Sanctuaries and the sacred in the ancient Greek world The First White Man Of The West Or The Life And Exploits Of Colonel Danl Boon, The First Settler Of Kentu Silver apples of the moon Elizabeth Jolley Tone it up nutrition plan regular Mauritanian Slavery Honda hr214 repair manual The Chaebol (Jae Bol) The church that takes on trouble My unlikely journey to ultramarathon greatness Forecasting with Judgment SKELETAL MUSCLE CONTRACTION The Jazz Pictures Federal-question jurisdiction and subject-matter jurisdiction generally An introduction to music in early childhood education Shakespeare and Gender: The Scholarly Literature (Shakespeare: The Critical Complex) Dont panic, Mother. Principles of pavement engineering Chronicles of the great rebellion. Uscg a school list Liars, extremists, and buaffalo wings Gods in the making and other writings