A Comparative Study on Statistical Software Packages with Reference to Graphical Tools

CHAYAN PAUL[1], HEMANTA SAIKIA[2], and DIBYOJYOTI BHATTACHARJEE[3]
1 University of Technology and Management, Shillong, Meghalaya, India
2 Kaziranga University, Jorhat, India
3 Assam University, Silchar, India

Software packages provide good platform to perform the statistical computations and data analysis. One of the important features of these packages is the tools available for drawing statistical graphs. Since graphs are considered to be important tool for analyzing data, often comparisons are made among the software packages based on the graphical tools available in them by default. This paper attempts to make a comparison among the statistical software packages based on the graphical tools available in these packages. Information about the graphical tools is collected from the various web sites. Ranks of the packages were calculated and a final list is prepared as per the ranks.

Keywords: *Statistical Graphics, ICT, Data Visualization.*

INTRODUCTION

During the late 50s and 60s three statistical software packages, BMDP, SPSS and SAS were developed. The first among them was BMDP and its development started in 1957, at the UCLA Health Computing Facility. Three Stanford University graduate students, Norman H. Nie, Dale H. Bent and C. Hadlai Hul, developed SPSS in the year 1968. SAS was developed almost simultaneously with SPSS, since 1968 by computational statisticians at North Carolina State University (Leeuw, 2009). With the development in the field of personal computer the second generation of statistical software started to appear in the market in the 80's. The two main statistical software packages that appeared in the market during the period were Data Desk in the year 1985 and JMP in 1989. Data Desk and JMP gave much emphasis on the graphical user interface as they were mainly developed for Macintosh platform. Features like dynamic graphics and graphical widgets to portray and interactively manipulate data sets were the main attractions in those software packages. While these two packages were famous for their graphical interfaces, second generation statistical software STATA was having command line interface. Development of STATA started in the year 1985. While the main features of Data Desk and JMP were graphical user interface (GUI) and interactive graphics, they were too difficult to extend. On the other hand, STATA mainly emphasized on extensibility and user generated code. In the early eighties, John Chambers and his team started developing S language, which was thought to be the alternative of MATLAB in Statistics. For history and development of S one can look into the works of Chambers (2008) and Becker (1994). Initially, S was distributed free of cost in academic institutions and was only used by the researchers in higher educational institutions. Insightful corporation later purchased S from AT & T and marketed it as S plus. S plus was fully dominating the market until R and LISP-STAT came in. In the year 1990, Luke Tierney developed LISP-STAT, a statistical environment embedded in the LISP interpreter. It became a good alternative to R, as it was more readily available and friendlier to personal computers. It became easier to extend the code written in either C or LISP. It had dynamic graphics capability and these graphics could be programmed and extended easily. During 2000, active development of LISP-STAT was stopped and R was available as an alternative. Ihaka and Gentlemen (1996) developed R incorporating some of the features from two earlier languages, S and Scheme.

In 1996, SPSS acquired BMDP and after that BMDP started disappearing from the market. Again in the year 2009, IBM acquired SPSS and changed its name to Predictive Analytics Software (PASW). As the name reflects, the focus of SPSS shifted from social science research to social science data analysis and business analytics. The same development is going on in SAS as well. Originally SAS was the abbreviated 'Statistical Analysis System'.

For history and development of Statistical software packages, one may go through the book by Francis (1981). He discussed about sixty software packages for statistical analysis. Hayes (1982) provided detailed study of various features and origin of 213 software packages available till 1980. Francis (1981) made the first systematic effort to measure the efficiency of the statistical software packages used in industry and academics. For history and development of statistical software one may see Foxwell (1984), an issue of Capital PC Monitor specially dedicated to statistical software packages available for IBM PCs. Other studies in this regard include Longley (1967), Wampler (1970), Wilkinson and Dallal (1977), Anscombe (1967), Hayes (1982), Wilkinson (1985), Simon and Lesage (1988, 1989), etc.

The developments in the field of Information and Communication Technology (ICT) take place in rapid speed. There are number of new software packages introduced into the market and old ones are modified several times to survive in this competitive market. Robertson and Nelson (2010) reviewed some software packages capable of space time disease surveillance analysis and analyzed some of their salient features, shortcomings and usefulness. Some more literature in this area includes Callert (2003), Oster (2003), Proctor (2006), Altman and Jackman (2011).

Because of the recent development of information and communication technology, collection and storage of large dataset has become easier. At the same time the complexity in analyzing these datasets also increased. Earlier when the size of the dataset used to be comparatively smaller, managing and analyzing the data were easy as well. If the result was to be reproduced, that too could be done without much effort. But, large datasets needs lots of calculations before they can be made ready for analysis and reproducing the result of analysis also got complex. With the near-exponential growth of PC computing power, many statistical techniques are available at the desk-top, provided by software packages that cover a wide range of analyses and statistical graphics (Morgan, 1998). A wide range of statistical software packages are available in the present market. The users often get perplexed to choose one package among them. Every software package seems to have its own set of unique features

A Comparative Study on Statistical Software
Packages with Reference to Graphical Tools

and ready to provide the user the best bid. Although the exact number of software packages available is hard to determine, but by a careful investigation through various websites, one can find existence of some several hundreds of software packages of these type. There are several websites providing such information, but specific mention can be made to the following web addresses,

http:/ / statpages.org/ javasta2.html

http:/ / www.math.yorku.ca/ SCS/ StatResource.html

http:/ / www.stata.com/ links/ statistical-software-providers/

http:/ / www.amstat.org/ careers/ statisticalsoftware.cfm

http:/ / en.wikipedia.org/ wiki/ Comparison_of_statistical_packages

These web addresses provide information about a large number of statistical software packages available in the current market. One can follow the links to visit the home pages of the software packages. Most of the packages are standalone, while a few are there which can be used as Add-Ins for various packages like Microsoft Excel. The software packages are mostly priced, while some of them are open sourced and can be downloaded free of cost. Most of the companies providing the priced software packages offer free trial versions for a limited period.

OBJECTIVES OF THE PAPER

The objective of the paper is to make a comparative study of the statistical software packages currently available in the market, based on the graphical tools available in those packages.

SELECTION OF THE SOFTWARE PACKAGES

The web addresses were visited in search of statistical software packages available. A total of 112 statistical software packages were found from the various sources. Table 1 shows the complete list in alphabetic order.

Table 1 List of Statistical Software Packages

Alphabet	Software Packages
A	ADaMSoft, Algebrator, Analyse-it, ASReml
B	Baudline, BMDP
C	CoStat and CoPlot - from CoHort Software
D	DADiSP, DAP, DataDesk, Dataplot, DataScene, Descartes (plotting tool)
E	EasyPlot, EditGrid, Epi Info, EpsTk, EViews
F	Fityk, FlexPro, FreeMat
G	GAUSS, GenStat, GeoGebra, GLE, GLIM - Genstat from the Numerical Algorithms Group (NAG), GNU Octave, Gnumeric, Gnuplot, Grace, GrADS, Graph, Graphis GraphPad Prism, Graphviz, Gretl
I	IDPS, IGOR Pro
J	JFreeChart, jHepWork, JMP
K	KChart, Kig, Kst
L	LabPlot, LabVIEW
M	Maple, Mapping Contouring System, MathCad, Mathematica, MATLAB, Mavis, Maxima, MedCalc, Minitab, MLPlot, Monarch Charts,
N	NCSS Statistical Software, NMath Stats, Nucalc, Numbers (iWork), NumXL
O	OpenEpi, OpenPlaG, Origin, OxMetrics
P	Paraview, PDL, PGPLOT, Physics Plot, ploticus, PLplot, Primer, PSPP
Q	QtiPlot
R	R, RATS, Revolution Analytics, RLPlot, RRDtool
S	SAGE, Salstat, SAS, SAS System, SciDAVis, Scilab SciPy, NumPy, matplotlib modules for Python, SHAZAM SigmaPlot, SigmaXL, S-Lang, SOCR, SPlus, S-PLUS, SPSS, Stata, Statgraphics, STATISTICA, Statistical Lab, StatPlanet, StatPlus, SymPy, Sysquake, Systat
T	Tableau, Teechart
U	UNISTAT
V	Visifire, VisIt
W	Winpepi, WPS
X	XploRe

After collecting names of the software packages, web sites of the respective software packages were visited for the purpose of collection of information regarding the availability of the data visualization tools in the packages. Some of the packages were not having any data visualization tools and hence those names were discarded from the list. Also, in case of some other packages, the websites were not having enough details about the data visualization tools in those packages. Hence, those were not considered further as well. After making a careful investigation, a list of 17 software packages were made from that list of 112. The rest were discarded mainly because either those were not having the sufficient data visualization tools in their packages or the websites were not having enough information regarding the availability of data visualization tools in the software packages. Table 2.2 lists the 17 software packages selected after filtering the list.

Table 2 List of software selected finally for comparison

Sl No	Name of the software
1	Analyse IT
2	BMDP
3	CoPlot
4	Data Plot
5	E View
6	Gauss
7	JMP
8	Mapple
9	Minitab
10	SAS
11	SPSS
12	STATA
13	Statgraphics Centurion
14	STATISTICA
15	SYSTAT
16	UNISTAT
17	Winks

METHODOLOGY

The information about the availability of the graphical tools in different software packages were collected visiting the websites. In some cases, where the information was not adequate in the respective websites, trial versions of those packages were installed and relevant information was collected. The whole list of data visualization tools were divided in three different categories, i.e., one dimensional, two dimensional and 3 or n-dimensional graphics. Then the frequency (number) of visualization tools for each packages under each category were collected. Appendix I shows the list of software and availability of the data visualization tools under above mentioned three categories. The table in Appendix I consider the name of software along the columns and the plots along the rows. The entry '1' in the cell indicates the presence of the corresponding plot in the software named along the column and '0' indicates its absence. The analysis can be performed in two steps.

Step I: Initially we perform Cochran's Q test to check if the different software have identical effects for the different graphical tools for the three different categories of plots viz. one dimensional, two dimensional and 3 or n-dimensional graphics separately. The Cochran's Q statistic (Cochran, 1950) is defined as,

$$Q = [K(K-1)]\left[\frac{\sum_{j=1}^{m}\left(C_j - \left[\frac{G}{K}\right]\right)^2}{\sum_{i=1}^{n} R_i(K-R_i)}\right] \sim \chi^2_{k-1} \qquad ...(1)$$

where K is the total number of software packages

C_j is the column total for the j^{th} software under a given category

R_i is the row total for the i^{th} graphical tools in a given category

G is the grand total for the category under consideration

A Comparative Study on Statistical Software
Packages with Reference to Graphical Tools

In case the test results indicates significant difference in the effect of the software then step 2 is taken up which is a type of composite index based on ranks.

Step 2: The software packages were having different frequencies under the three categories. If a single category is given importance and ranks are calculated depending on this, the calculated rank may not reflect the importance of the other two categories. Thus, in the next step of calculation, the ranking, weight for each of the categories were calculated. Generally the simple average gives equal importance to each of the categories, but when variables are weighted to a composite measure, the relative importance of the variables is considered. Iyenger and Sudarshan (1982) opined that the weights vary inversely as the variance in the respective variables. This definition of weight has been used to calculate the weights for the three categories of the graphical tools.

Let x_{ij} be the frequency of graphical tools in the i^{th} category for j^{th} software, where i = 1, represents one dimensional graphical tools, i = 2, represents two dimensional graphical tools and i = 3, represents 3 or n-dimensional graphical tools and j (= 1,2,…, 17) represents the different software packages in the list. If wi be the weight of i^{th} category then it is given by

$$w_i = \frac{c}{\sqrt{\mathrm{var}(x_i)}}, i = 1,2,3 \qquad ...(2)$$

Where $\sum w_i = 1$ and C is a normalizing constant which follows $c = \left[\sum_{1}^{3} \frac{1}{\sqrt{\mathrm{var}(x_i)}}\right]^{-1}$

The choice of the weights in this way would ensure that the large variation in any one of the factor would not unduly dominate the contribution of the rest of the factors (Iyenger and Sudarshan, 1982).

Following this we calculate a score S_j (j=1,2,…17) for each software using the formula in (3).

$$s_j = \sum_{i=1}^{3} w_i * x_{ij} \qquad ...(3)$$

The next step is to calculate the ranks for the different software packages and for that a well-established method by Olson (2004) was used. The steps of calculation in this method are stated below:

a) Identify the ideal alternative (extreme performance on each criterion) s+:

b) Identify the nadir alternative (reverse extreme performance on each criterion) s-:

c) Develop a distance measure over each criterion to both ideal (D+) and nadir (D-).

d) For each alternative, determine a ratio R equal to the distance to the nadir divided by the sum of the distance to the nadir and the distance to the ideal,

$$R = \frac{D^-}{D^- + D^+} \qquad ...(4)$$

e) Rank order alternatives by maximizing the ratio in Step d.

Following the above mentioned steps , the researchers calculated the ideal alternative of S_j i.e., maximum of S_j and nadir alternative of S_j i.e., minimum of S_j. Then the authors calculated the differences D+, the differences of each S_j from maximum of S_j and D-, differences of each S_j from minimum of S_j. The ranks for each software packages are calculated by using formula in (4).

RESULTS

Table 3. shows the result of the Cochran's Q test. The performance of the packages under different categories of plots viz. one dimensional, two dimensional and 3 or n-dimensional graphics are not same indicated by p-value (less than 0.05) for each of the cases.

Table 3: Result of Cochran's Q test on data in Appendix			
	Calculated value of Cochran's Q-Statistic	p-value of $\chi^2_{0.05}$ for 16 df	Conclusion
1Dimensional	33.643	0.006	Difference is significant
2Dimensional	41.798	0.0004	Difference is significant
3D/N dimensional	34.949	0.004	Difference is significant

As the effect of the different software varies significantly in each of the different categories so a composite index is developed based on the procedure explained in Step 2 of the previous section. Table 4 provides list of 17 software packages along with their composite score and calculated ranks. The calculations of the ranks are shown in the Appendix II.

Table 4 Calculated scores and ranks of the software packages depending on the available graphic features		
Software	**Score**	**Rank**
Mapple	1.0000	1
STATISTICA	0.9252	2
Statgraphics Centurion	0.8826	3
SPSS	0.8772	4
JMP	0.8077	5
UNISTAT	0.6065	6
Minitab	0.5962	7
SAS	0.5728	8
Data Plot	0.4815	9
E View	0.4712	10
Analyse IT	0.3949	11
STATA	0.3729	12
CoPlot	0.3716	13
SYSTAT	0.1703	14
Gauss	0.1511	15
BMDP	0.0220	16
Winks	0.0000	17

CONCLUSION

The ranks calculated for the 17 software packages here are totally based on the graphical tools available in those packages. From Table 4, one can observe that relatively lesser known software package 'Mapple' secured the top position amongst the 17 packages considered for calculation. One of the most popular software used in statistical analysis, SPSS (now IBM SPSS) could only secure the 4th position in the list preceded by Mapple, Statistica and Statgraphics Centurion. SPSS is mainly used in the field of social science research and business analytics. SAS is also in the same field of business analytics and it needs its user to write codes to get the computation done. Minitab commonly used for teaching statistics to the students. The ranks calculated here are only based on the graphical tools and it may change if other analytical features of the software are considered.

A Comparative Study on Statistical Software
Packages with Reference to Graphical Tools

Appendix I

	Analyse IT	CoPlot	BMDP	Data Plot	E View	Gauss	JMP	Mapple	Minitab	SAS	STATA	Statgraphics Centurion	STATISTICA	SYSTAT	UNISTAT	Winks	SPSS
One dimensional																	
Bar Plot	1	0	1	0	1	1	1	1	0	1	1	1	1	1	1	1	1
Histogram	1	1	0	1	0	1	1	1	1	0	1	1	1	0	0	0	1
Boxplot	1	0	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1
Mean Error Bar Plot	1	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	1
XY Graph / Contour Plot	0	1	0	1	0	1	1	1	1	1	0	1	1	0	1	0	0
Error Plot	0	1	0	0	1	0	0	1	0	0	0	0	1	0	1	0	1
Stacked Bar Plot	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1
Block Plot	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0
Mean Plot	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	1
Normal Probability Plot	0	0	0	1	0	0	1	1	1	0	0	1	1	0	0	0	1
Hi-Lo-Open-Close chart	0	0	0	0	1	0	0	0	0	0	0	1	0	0	1	0	1
Two dimensional																	
CDF Plot	1	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0
Scatter Plot	1	0	0	1	1	0	1	1	1	1	1	0	1	1	1	0	1
Normal QQ Plot	1	0	0	0	0	0	1	1	0	0	0	1	1	0	0	0	1
Frequency Plot	1	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0
Polar Graphs	0	1	0	0	0	1	0	1	0	0	0	0	0	0	1	0	1
Bi Histogram	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	1
Control Chart	0	0	0	1	0	0	1	1	1	0	0	1	0	0	0	0	1
SD Plot	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	1
Area Plot	0	0	0	0	1	0	1	1	0	1	1	0	0	1	1	0	1
Spike Plot	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0
Seasonal Plot	0	0	0	0	1	0	0	1	1		0	0	0	0	0	0	1
Bland Altaman Plot	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0
Pareto Curve	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1
Density Chart	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0
Three/n dimensional																	
Scatterplot Matrix / correlation matrix	1	1	0	0	0	0	1	0	1	0	1	1	1	0	0	0	1
3D Graph	0	1	0	0	0	0	0	1	1	0	0	1	1	0	0	0	1
Triangle Graph	0	1	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0
Star plot	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0
3D Scatter Plot	0	0	0	0	0	0	1	0	0	0	0	0	1	0	1	0	1
Bubble Plot	0	0	0	0	0	0	1	1	0	1	0	1	0	0	0	0	0
Surface Plot	0	0	0	0	0	0	0	1	0	1	0	1	1	0	1	0	0
Radar Chart	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	
Icon Plot	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0

Appendix II

Software	Uni Dimensional	Two Dimensional	Thee/n Dimensional	Sij	diff with maximum	diff with minimum	D+	D-	R	Rank
Analyse IT	4	5	1	3.01522	-3.34969	2.1863366	3.34969	2.186337	0.394929	11
CoPlot	4	1	3	2.886024	-3.478886	2.0571414	3.478886	2.057141	0.371592	13
BMDP	2	1	0	0.950622	-5.414288	0.121739	5.414288	0.121739	0.02199	16
Data Plot	6	4	1	3.494412	-2.870498	2.6655293	2.870498	2.665529	0.481488	9
E View	5	7	0	3.437271	-2.927639	2.6083879	2.927639	2.608388	0.471166	10
Gauss	4	1	0	1.665529	-4.699381	0.8366463	4.699381	0.836646	0.151128	15
JMP	5	8	4	5.300312	-1.064598	4.4714293	1.064598	4.471429	0.807696	5
Mapple	6	11	4	6.36491	0	5.5360269	0	5.536027	1	1
Minitab	4	8	2	4.129195	-2.235715	3.3003123	2.235715	3.300312	0.596152	7
SAS	4	4	4	4	-2.36491	3.171117	2.36491	3.171117	0.572815	8
STATA	3	6	1	2.893481	-3.471429	2.0645976	3.471429	2.064598	0.372939	12
Statgraphics Centurion	7	5	5	5.714907	-0.650003	4.8860243	0.650003	4.886024	0.882587	3
STATISTICA	7	6	5	5.950622	-0.414288	5.121739	0.414288	5.121739	0.925165	2
SYSTAT	1	6	0	1.771742	-4.593168	0.9428586	4.593168	0.942859	0.170313	14
UNISTAT	5	5	3	4.186337	-2.178573	3.3574537	2.178573	3.357454	0.606474	6
Winks	1	2	0	0.828883	-5.536027	0	5.536027	0	0	17
SPSS	9	12	3	7.266154	0.901244	6.4372709	0.901244	6.437271	0.87719	4
SD	2.124783726	3.222165877	1.866894244							
1/SD	0.470636135	0.310350254	0.535648981							
K	0.759511724									
Weight	0.357453662	0.235714657	0.406831681							
max (Sij)	6.364909921									
min (Sij)	0.828882976									

REFERENCES

BMDP Statistical Software. (2012). Retrieved May 28, 2013, from http://www.statistical-solutions-software.com/: http://www.statistical-solutions-software.com/bmdp-statistical-software/bmdp/

The GAUSS™ Mathematical and Statistical System. (2012). Retrieved May 29, 2013, from http://www.aptech.com/: http://www.aptech.com/products/gauss-mathematical-and-statistical-system/

About IHS EViews. (2013). Retrieved May 29, 2013, from http://www.eviews.com/: http://www.eviews.com/general/about_us.html.

Analyse It Standard Edition. (2013). Retrieved May 28, 2013, from http://analyse-it.com/: http://analyse-it.com/products/standard/

JMP® STATISTICAL DISCOVERY SOFTWARE. (2013). Retrieved May 29, 2013, from http://www.jmp.com/: http://www.jmp.com/software/jmp10/

Altman, M., & Jackman, S. (2011). Nineteen Ways of Looking at Statistical Software. Journal of Statistical Software, 1-11.

Anscombe, F. (1967). Computing in Statistical Science through APL. New York: Springer-Verlag.

Becker, R. (1994). A Brief History of S. Technical report. Retrieved March 6, 2013, from http://www2.research.att:http://www2.research.att.com/areas/stat/doc/94.11.ps.

Callaert, H. (2003). Comparing Statistical Software Packages: The Case of the Logrank Test in StatXact. The American Statistician, 214-217.

Chambers, J. M. (2008). Software for Data Analysis Programming with R. California: Springer.

Cochran, W.G. (1950). The Comparison of Percentages in Matched Samples. Biometrik, 37(3/4), 256–266.

CoPlot Software for Great Scientific Graphs, Maps, and Technical Drawings. (n.d.). Retrieved May 28, 2013, from http://www.cohort.com/: http://www.cohort.com/coplot.html

Foxwell, H. J. (1984). Capital PC Monitor. Gaithersburg.: Capital PC User Group.

Francis, I. (1981). Statistical Software: A comparative Review. New York.

Hayes, A. (1982). Statistical Software: A Survey and Critique of its developement. Arlington: Office of the Naval Research.

Heckert , A., & Filliben, J. (2002, October 9). Dataplot summary. Retrieved May 28, 2013, from http://www.itl.nist.gov/: http://www.itl.nist.gov/div898/software/dataplot/summary.htm

Ihaka, R., & Gentleman, R. (1996). R: A Language for Data Analysis and Graphics. Journal of Computational and Graphical Statistics, 299-314.

Leeuw, J. D. (2009, December 23). Statistical Software - Overview. Retrieved February 12, 2013, from http://statistics.ucla.edu: http://statistics.ucla.edu/system/resources/BAhbBlsHOgZmSSI+MjAxMi8wNS8xNi8xNF81MV8yNF8xNzhfU3RhdGlzdGljYWxfU29mdHdhcmVfT3ZlcnZpZXcucGRmBjoGRVQ/Statistical%20Software%20-%20Overview.pdf.

Longley, J. W. (1967). An appraisal of least-squares for the electronic computer from the point of view of the user. Journal of the American Statistical Association, 819-841.

Morgan, W. T. (1998, February). A Review of Eight Statistics Software Packages for General Use. The American Statistician, 52, 70-82.

Oster, R. A. (2003). An Examination of Statistical Software Packages for Categorical Data Analysis Using Exact. The American Statistician, 201-213.

Oslon, D. L. (2004). Comparison of Weights in TOPSIS Models. Mathematical and Computer Modelling, 721-727.

Proctor, J. L. (2006). SPSS vs. excel: Computing software, criminal justice students, and statistics. Journal of Criminal Justice Education, 433-442.

Robertson, C., & Nelson, T. A. (2010). Review of software for space-time disease surveillance. International Journal of Health Geographics, 9-16.

Simon, S. D., & Lesage, J. P. (1988). Benchmarking numerical accuracy of statistical algorithms. Computational Statistics and Data Analysis, 197¬-209.

Simon, S. D., & Lesage, J. P. (325--332). Assessing the accuracy of ANOVA calculations in statistical software. Computational Statistics and Data Analysis, 1989.

Tanenbum, A. S., & Woodhull, S. A. (2001). Operating Systems Design and Implementations. New Delhi: Prentice Hall of India.

Tierney, L. (1990). LISP-STAT: An Object-Oriented Environment for Statistical Computing and Dynamic Graphics. New York: John Wiley and Sons.

Wampler, R. H. (1970). A report on the accuracy of some widely used least squares computer programs. Journal of the American Statistical Association, 549-565.

Wilkinson, L. (1985). Statistics Quiz. Evanston: SYSTAT, Inc.

Wilkinson, L., & Dallal, G. E. (1977). Accuracy of sample moments calculations among widely used statistical programs. The American Statistician, 128-131.

BIOGRAPHIES

DibyojyotiBhattacharjee, Ph.D. is an Associate Professor in the Department of Business Administration, Assam University, Silchar, Assam, India. He has an M. Sc and an M. Phil in Statistics from University of Delhi. He pursued his Ph. D in the area of Statistical Graphics from Gauhati University. He has to his credit around 80 research publications in various national and international journals. ". His research papers have been published in reputed journals like Vikalpa, International Journal of Performance Analysis in Sports, Demography India etc. He has also authored/co-authored 15 books in different subfields of statistics. He has to his credit almost 15 years of teaching experience. He served institutions like GC College, Ramlal Anand College, Central Statistical Organization, Gauhati University etc. He is a member of editorial board of several national and international journals.

Hemanta Saikia, Ph.D. is an Assistant Professor of Statistics in the School of Business, Kaziranga University, Jorhat (Assam). He did his M.Sc. in Statistics with specialization in Data Mining & Computational Mathematics from Gauhati University. He holds a Ph. D degree in the area of "Data Mining in Sports" from Assam University, Silchar. His research papers have been published in reputed journals like Vikalpa: A Journal of Decision Makers (IIMA), Sports Business and Management: An International Journal (Emerald), International Journal of Sports Science and Coaching (Multi-Science Publishers), International Journal of Sports Science and Engineering (Academic Press), etc. He is a life member of Indian Association of Social Science and Health (IASSH) and North Eastern Economic Association (NEEA). He has also presented papers in various national as well as international conferences across India.

Chayan Paul, Ph.D. is an Assistant Professor in the Department of Computer Science, University of Technology and Management, Shillong. He is an M. Sc in Computer Science from Assam University. He also completed his Ph.D. in the area of Data Visualization from Assam University, SIlchar. He has published papers in national and international journals. He has also presented papers in various national and international conferences.