Underpinning Statistical Computing Knowledge and Skills for Students

LI Wing Kin, Ken

Department of Information and Communications Technology, Hong Kong Institute of Vocational Education (Tsing Yi)

Some software developers are devoted to designing and implementing computer programs to offer a window-based environment, handy statistical tools and user-friendly features, such as pull-down manuals, an on-line Help menu, hyperlinks to statistical glossary, fancy computer output and eye-catching graphical displays without taking into account whether such tools can serve statistical purposes (Li, 2002). Statistical software does not address non-statistical questions inherent in a statistical problem. Addressing these questions is beyond its capability but is left to users of statistical software to attempt (Chambers, 2000). In addition, teaching how to program statistical software, however, is over-emphasised by some lecturers and students may not develop the skills needed to be competent at justifying and/or interpreting statistical results (Li, 2005). As such, the author of this paper advocates a model of teaching statistical software to reinforce students' understanding of how to use statistical software properly and efficiently.The model so developed follows the process of statistical analysis of data and consists of six major steps: examining data characteristics, selecting statistical tools, understanding the strengths and weaknesses of statistical software, checking the accuracy of statistical output, interpreting statistical output and presenting statistical results. First, examining data characteristics is to conduct a preliminary study on the characteristics of given data in terms of the types, format, content, context, measurement and measurement units of data. Second, correct selection of statistical tools should involve checking whether or not an interpretation of statistical results derived from a specific statistical tool reflects the contextual meaning. Third, understanding the strengths and weaknesses of statistical software refers to the awareness of capability, limitations as well as computational quality of the software a data analyst will use. Fourth, checking the accuracy of statistical output is to assure an absolutely correct or accurate statistical output in terms of statistical logic and reasoning. Fifth, interpreting statistical output is to interpret data as evidence (Royall, 1999). Sixth, presenting statistical results is about the dissemination of statistical results through written reports which is a vital component of statistical work. A questionnaire-based survey was then conducted to study how useful Higher Diploma in Business Analysis students found the teaching model. The results of the survey indicated that many of the students held positive perceptions of adopting the six-step model.

References :

Chambers, J.M. (2000). Users, programmers, and statistical software. *American Statistical Association Journal of Computational and Graphical Statistics, 9*(3), 404-422.

Li, K.W. (2005). Pedagogical concerns of SAS teaching. *Proceedings of SAS Academic User Group Conference 2005*, pp. 22-29.

Li, K.W. (2002). Enhancing students' graphic communication. *Proceedings of Science and Technology Education Conference 2002*, pp. 414-422.

Royall, R. (1999). *Statistical Evidence*. New York: Chapman & Hall/CRC.