

Integrating MS-Excel in Research Methodology Course

Rohit Vishal Kumar

Department of Marketing

Xavier Institute of Social Service, Ranchi

1. Introduction

I joined academics in the year 2001 after serving with marketing research companies in various cities in India. As is wont to happen with a junior academician, I was asked to teach Business Statistics & Research Methodology to the students. The teaching of the course was, though not the problem; the problem started when the students went for their summer training and subsequently started calling me up asking “how to analyse the data?” My typical response was to use SPSS. However, I soon realized that neither all students had access to the software nor had the Institute a site license for the software. In effect, by asking students to use SPSS I was indirectly contributing to software piracy. Secondly, without adequate and proper knowledge of the software, the students were coming up with “*wild analysis*” – like calculating Pearsonian correlation for nominal data, finding chi-square for percentage or ratio data etc.

Reflecting on the problem, I started consulting colleagues in other institutes on an informal basis. I soon realized that the problem was widespread. These consultations revealed three major problems (i) Second and Third level institutes did not invest in analytical software for student consumption primarily because of the high cost of the software (ii) Even if the B-School had the software it did not have competent resource person for imparting software knowledge and (iii) nowhere was the aspect of “Data Entry & Data Coding” taught to the students. In a bid to overcome these problems, I started looking for a solution and finally decided to integrate Microsoft Excel in the syllabus of Research Methodology.

2. Statistical Features of MS-Excel

The choice of using MS-Excel was due to the fact that most b-schools had purchased computers to set up their computer lab and these machines already had the OEM version of MS-Office installed on them. Microsoft Excel, first released in 1985 has grown into a comprehensive spreadsheet which can also double up as an analysis software, charting software, flat-form database and a full fledged programming environment (Wikipedia, 2009).

The MS Excel 2003 version contains about 80 statistical functions. Besides allowing for calculation of descriptive statistics like measures of central tendency, dispersion, skewness and kurtosis, it also has functions to compute correlation, linear regression slope and intercept. It even allows testing of hypothesis by providing functions like chi-test, f-test, t-test and z-test. It has functions to calculate the probability distributions of the major distributions – beta, binomial, chi-square, exponential, f, gamma, hyper-geometric, log normal, negative binomial, normal, poisson, t and weibull.

Another powerful feature is the provision of “array formulas”. An array formula can perform multiple calculations and then return either a single result or multiple results. The usage of array formula’s acts as a boon during analysis – especially when trying to tabulate raw data into frequency classes. MS Excel also provides “Data Analysis Tool Pack” which is available as an add-on in all the versions of the software. This tool pack acts as a graphical user interface between the user and the data and provides some additional analysis facilities to the user. The additional analysis that can be done using the tool pack include – ANOVA, Exponential Smoothing, Fourier Analysis, Histogram, Moving Average, Rank and Percentile, Multiple Regression and Sampling. Besides it allows quick calculation of Descriptive Analysis and of Z and T tests for hypothesis testing. The charting facility given by MS-Excel allows creation of basic charts and graphs along with insertion of trend lines and error bars in almost all of the charts. It allows the following trend lines – linear, logarithmic, polynomial, power exponential and moving average – along with the adjusted R^2 values. Pivot Table’s and Pivot Charts are another powerful features of MS Excel which allow us to generate 2x2 tables and charts to study relationships in an interactive way.

3. Problems in Integrating MS-Excel with Research Methodology

Given the range of features that are present in MS-Excel, it was a very good choice for statistical data analysis; however I faced many problems in teaching the students.

First and foremost, most of the students were ill-aware of the features of MS-Excel and especially of the types of referencing – absolute, mixed or relative – which is used by the software for recalculation purpose. So in a way I had to teach the students the basics of MS-Excel from the beginning. Secondly, the students were unaware of the principles of data coding and data entry. For this purpose, I had to conduct special classes devoted specially to the subject. Thirdly there was impatience amongst students regarding the amount of time required to enter data in the software. Most students were given an average of 200 samples (sometimes even more) by the companies and assuming that a questionnaire had 50 data entry columns; that required entering data for 10,000 cells in the software. It was a job which required patience but the students wanted a quick-fix solution. Fourthly, I had to counsel the students regarding drawing up an “analysis plan” before actually analyzing the data. This also was considered as an irk-some process by most of the students.

However, the biggest problem which I faced was in trying to integrate the theory of MS-Excel software along with Research Methodology paper. Students considered MS-Excel to be a part of “Computer Applications” paper and were disgusted to study it once again as an adjunct to the “Research Methodology” paper. I overcame this problem, by introducing MS-Excel from the initial “Business Statistics” paper. I would show them how to calculate mean, median, etc. on the computer while the subject was being taught. This generated interest amongst students regarding the use of the software and the next batch was more willing to integrate MS-Excel with research methodology.

4. Conclusion

It has been more than five years since I am following this approach and I would say that it has helped the students to gain a better insight in research methodology. Opponents of MS-Excel do point out that there are various lacunae in the software (Burns, 2009; Cox, 2000; Goldwater, 2007; McCullough & Heiser, 2008). Even though I personally agree with these criticisms, but until and unless statistical software companies start selling statistical packages at a cost affordable to average Indian student, MS-Excel seems to be the best option available for data analysis. I strongly recommend that business-schools in India should start laying extra emphasis on the use of MS-Excel or its freeware alternatives like Gnumeric or OpenOffice for data analysis.

References:

Burns, P. (2009). Spreadsheet Addiction. Retrieved September 15, 2009, from http://www.burns-stat.com/pages/Tutor/spreadsheet_addiction.html

Use of Excel for Statistical Analysis. <http://www.agresearch.co.nz/Science/Statistics/exceluse1.htm>

Goldwater, E. (2007). Using Excel for Statistical Data Analysis - Caveats. Retrieved September 15, 2009, from <http://www-unix.oit.umass.edu/~evagold/excel.html>

McCullough, B. D., & Heiser, a. A. (2008). On the accuracy of statistical procedures in Microsoft Excel 2007. *Computational Statistics & Data Analysis*, 52(10), 4570-4578.

Wikipedia. (2009, September 15). Microsoft Excel. Retrieved September 15, 2009, from http://en.wikipedia.org/w/index.php?title=Microsoft_Excel&oldid=314154110