Chapter 7
Excel-lence in Data Visualization?
The Use of Microsoft Excel for Data Visualization and the Analysis of Big Data

Jacques Raubenheimer
University of the Free State, South Africa & The University of Sydney, Australia

ABSTRACT

Spreadsheets were arguably the first information calculation and analysis tools employed by microcomputer users, and today are arguably ubiquitously used for information calculation and analysis. The fluctuating fortunes of PC makers coincided with those of spreadsheet applications, although the last two decades have seen the dominance of Microsoft Excel in the spreadsheet market. This chapter plots the historical development of spreadsheets in general, and Excel in particular, highlighting how new features have allowed for new forms of data analysis in the spreadsheet environment. Microsoft has undoubtedly cast Excel as a tool for the analysis of big data through the addition and development of features aimed at reporting on data too large for a spreadsheet. This chapter discusses Excel's ability to handle these data by means of an applied example. Data visualization by means of charts and dashboards is discussed as a common strategy for dealing with large volumes of data.

DOI: 10.4018/978-1-5225-2512-7.ch007

INTRODUCTION

Since the dawn of writing, humans have endeavored to capture data. Early tools were primitive, with limited capacity (as with early clay tablets inscribed with a stylus) and limited durability (as papyrus and vellum degraded quickly over time). None can argue that both storage capacity and longevity have increased since the introduction of electronic computing and electronic storage, paralleling what has come to be known as Moore's law (cf. Moore, 1965) even though Moore himself notes that his "law" will eventually prove to be unsustainable (Courtland, 2015). Martin Hilbert (2012; cf. also Hilbert & López, 2012a, 2012b) defines the start of the digital age as that point at which the world's digital storage capacity exceeded its analog storage capacity—which he puts at the year 2002. It is exactly this growth in storage capacity that has facilitated the rise of big data. But with storage comes the desire to retrieve and then to analyze, and for this, we need tools that can handle big data.

Microsoft Excel was a relatively late entrant to the spreadsheet market, but since then it has become the de facto spreadsheet program, holding—by some estimates—up to 95% of the market share for spreadsheets. As such, Excel is, for many, the first line tool they turn to for data analysis and data visualization. For better or worse, Microsoft has placed voluminous resources behind its efforts to position Excel as a big data analysis tool (Sharma & Ellis, 2016). This chapter will plot that development and will examine what Excel offers the user who wants to make forays into working with big data.

HISTORY

The democratization of electronic information can truly be said to have started with the advent of the personal computer, and spreadsheets have played a central role in that process, to such an extent that Hesse and Scerno (2009) pointedly state:

It was VisiCalc (in approximately 1981) and the development of electronic spreadsheets that advanced PC sales beyond word processing machines and changed the definition of secretary (eliminating much of the work that secretaries had done) to administrative assistant. Wang redefined the secretary's job; however, with their word processing and spreadsheet capabilities, PCs redefined the secretary…. We contend that without the electronic spreadsheet, PCs might have remained hobby tools or game machines. (pp. 159-160)