Chapter 61
Big Data Analytics in Higher Education

Tanveer H. Shah
University of Azad Jammu and Kashmir, Pakistan

ABSTRACT

This chapter reviews the literature on the use of business analytics in higher education. Universities have large datasets available to predict future direction and generate actionable information. An important type of analytics used to improve management processes and to make informed decisions is big data business analytics. State university executive leaders may improve the effectiveness of their decisions by integrating business analytics in the decision-making models. However, there is a need to examine the use of big data business analytics in the decision-making process at the executive leadership level of the selected state universities. Especially in the context of how descriptive, predictive, prescriptive, decisive and basic analytics, and data collection influence the decision-making process at the executive leadership level of the state universities in terms of student retention and graduation rates.

INTRODUCTION

The range of challenges that state universities face today is vast and continuously expanding. These challenges are political, economic, societal, and cultural. Higher educational leaders are struggling to meet these challenges and compete in a highly competitive environment. Universities have taken several initiatives to meet these challenges. One of these initiatives is the use of analytics to make informed decisions at the executive leadership level. Universities have large datasets available to predict future direction and generate actionable information. The integration of big data business analytics in decision-making processes at the executive leadership level of state universities may result in their increased effectiveness. Delen and Demirkan (2013) stated that big data analytics offer a great potential to provide much needed information for effective decision-making. In order to use big datasets and analytics effectively, it is important to understand what is big data and what is analytical process.

DOI: 10.4018/978-1-6684-3662-2.ch061

BACKGROUND

The term big data referred to large datasets first coined in the 1990s. The definition of these large datasets had been changing over time. In the 1980s, this term referred to a size of data that was so large that magnetic tapes were required to store it. In the 1990s, big data referred to any size of data that was beyond the limitations of Microsoft Excel and desktop PC. Currently, big data refers to any data that is too large to place and analyze in a desktop relational database and such data require specialized software and computing techniques (Jacobs, 2009).

Diebold et al. (2012) claimed that big data emerged from statistics, econometrics, and computer science disciplines. Big data utilized knowledge, skills, and techniques of these disciplines in a coordinated manner. John Mashey at a lunch table in Silicon Graphics Inc. coined the term big data and elaborated on its use. Diebold et al. (2012) further claimed that Mashey, Weiss and Indurkhya, Diebold, and Laney were among the earlier contributors to the term big data. Diebold concluded that big data had emerged as a discipline of its own.

Buhl (2013) indicated that there was a lot more to come in the field of big data such as sensor-generated data. The term big data did not only refer to traditional data but it also included the analysis of huge amounts of data. Analysis provided new opportunities to improve the management and decision-making processes. Big data provided more options for decision-making, but it also added complexity to this process. Therefore, besides knowing big data itself, knowing how to use big data was equally important (Buhl, 2013).

Buhl, Röglinger, Moser, and Heidemann (2013) claimed that big data was multidiscipline and evolved from the combination of emerging technologies of data storage and data processing. The characteristics of big data include volume, velocity, variety, and veracity of data. Veracity meant maintaining the quality of data.

According to Davenport, Barth, and Bean (2013), in big data, data flows were more important than data stocks. In a big data environment, quick analysis, decisions, and actions were required. It was important to keep pace with the emerging trends in data flows that kept changing continuously. Observing trends was useless unless leaders took specific actions identified by recognizable trends (Davenport et al., 2013).

Davenport further explained that big data environment needed data scientists instead of data analysts. These professionals possessed multidisciplinary skills of information technology, analytics, social network sociology, statistics, communication, and programming skills (Davenport et al., 2013).

In the journal *Strategic Direction*, the publishing director Tony Roche (2013) pointed out important differences between big data and traditional data. It was not only size but also the form of data that mattered. Big data was not just storage of historical data like traditional data rather it was constant, current, live, dynamic data that flew through an organization. It forced the organization to change information management systems in order to take advantage of these qualities of big data. Big data was not meant to be buried within the computers. Rather, it had to go to decision-makers who could use it to act quickly. The process of using big data would make the core functions of organizations appear completely different (Direction, 2013).

Schermann, Hemsen, Buchmüller, Bitter, Krcmar, and Markl (2014) elaborated the big data definition further. According to their study, big data referred to technologies used to provide right information to the right persons in the right volume and the right quality at the right time.