Data Science and Organizations

Marc S. Galli

Walden University

Professor Dr. Mike Collins

DSCI 2010: Data Science Essentials

July 18, 2020

Data Science and Organizations

Science is as much a practical activity as it is an intellectual one and involves systematic study through observation and experiment (Lexico, 2020). The lexiconic term science is multidisciplinary in its most generic sense, yet, specificity may be acquired through conjunction with a chosen subject matter. Just as biological science is the study of our biosphere, data science is essentially the study of data (IBM, 2016). Philosophically, data collectively consists of things known or assumed to be facts which constitute the basis of reasoning or calculation (Oxford, 2020). Datum, individual facts and statistics, when collected together, are then used for reference and analysis. Ergo, in more finite terms, data science encompasses the study of structured and unstructured data as we seek to understand and to discover insights and trends in an effort to make more strategic corporate decisions and informed choices (IBM, 2016).

Data science has become increasingly more critical to businesses small and large as managers unilaterally seek to interpret and comprehend singularities to the end that we desire to see them in a greater context and recognize correlations to ascertain a grander meaning (International Federation of Classification Societies, 1998). Due, in large part, to the ubiquity of the internet and the interconnectedness of computers and people today, data is more available than it has ever been. A day in age existed when people used to purchase with cash. This day and system has succumbed to the ease of transaction and use of debit and credit cards as well as virtual and online banking, which has served to enlarge the subset of usable data. Corporations today can gather statistics and metrics on what companies' websites we visit, what purchases we make, and what products we look at and buy. So much data collection exists that even the number of seconds one spends viewing a product online can be captured. Software that once used to be proprietary and expensive is now open source and free. The additional impossibility of containing the data collected is now possible since storage is so much less expensive than in days past (Big Data University, 2016). For these reasons, corporate America now possesses the capability and the means necessary to study their customers and even the activity of potential clientele, and with that comes the desire to interpret the mass of data collected and make predictions on behavior to seek to better convert guests into paying customers.

Big data is a term to describe a massive amount of data, which is so large that it would be nearly impossible to process the data via traditional means. If the scope of data which a small company can collect can be comprehended by comparing that data subset to a slow trickle coming from a leaky faucet, then the data subset known terminologically as big data can be equated to the volume of water exiting a fire hydrant (Kiran, 2019). Yet, big data is more available and accessible than ever, even to smaller companies and organizations thanks to open-source tools, cloud computing, and data visualization methods (Big Data University, 2016). To elaborate, proprietary software has been mimicked or reverse engineered so that it has been made available to the public as open-source. Now, any kid in his garage or individual with some programming background and knowledge can improve upon and build on the data analysis methods once reserved for credentialed data scientists. The software he needs is open-source and free, and computing power can be leveraged in the cloud. Data science is multidisciplinary and more methods of visualizing data exist today than ever before.

As a case study, insights gained through data science enable Nutri Mondo to not only aid local communities in gaining better access to healthy food but also to address the injustices which lead to hunger on a global scale. The value of data scientists' teamwork cannot go understated. In the case study, Nutri Mondo capitalized on a team of four, to wit: a research manager, a data analyst, a project manager, and a director for education development and outreach. Collectively the team understood the problem, which is that, in spite of a global overabundance of food, millions still remain underfed (Laureate, 2016). Additional consideration has been given for health issues that surround this sustenance imbalance, and a solution was derived through data science's enhanced study and analysis. The charted solution involved enabling and enhancing methods for more food distribution.

In conclusion, in this day and age, data has never been more ubiquitous and methods for capturing and cultivating this data have never been more possible. Through the field of study known as data science, analysis is possible by even small businesses and organizations. Intellectuals today are less and less interested in proprietary methods and technologies and are more interested in leveraging the collective intelligence of the masses and building on each others' work for the furtherance of societies local and abroad.

References:

Big Data University. (2016). *Why data science is relevant* [Video file]. Armonk, NY: Author.

IBM. (2016). *What is data science?* [Video file]. Armonk, NY: Author.

International Federation of Classification Societies. Conference. (1998). *Data Science,*

Classification, and Related Methods: Proceedings of the Fifth Conference of the

International Federation of Classification Societies (IFCS-96), Kobe, Japan, March 27–

30, 1996. Springer.

Kiran, R. (2019, December 20). *What Are The Characteristics of Big*

Data? Edureka. https://www.edureka.co/blog/big-data-characteristics/

Laureate. (2016). *Nutri Mondo*. Laureate International Universities. https://mym.cdn.laureate-

media.com/2dett4d/IBM/LIDS0000/EN/nutrimondo/index.html

Lexico. (2020). *Science*. Oxford. https://www.lexico.com/en/definition/science

Oxford. (2020). *Data*. Lexico. https://www.lexico.com/en/definition/data