

INCORPORATING ANALYTICS HOLISTICALLY FOR BOTH BUSINESS & SOCIETY

Anushka Jain

This course has assisted one in becoming a well-rounded analytical professional capable of translating data into strategic decision making. The Kelsey-White simulation introduced the concepts such as business analytics, decision making, business intelligence, marketing research, and marketing strategy. It emphasized on both technical and managerial aspects of business analytics, allowing one to make better data-driven decisions and solve problems more quickly. It emphasized on using historical data to gain insights, understand the future by improving forecasting and efficiency while determining the best course of action. Thus, it is necessary to integrate descriptive, diagnostic, and predictive analytics with business acumen concepts such as financial, strategic, and business goals (Caruso, 2018). As a result, one must answer the following questions: "What happened, and what might happen? and finally, what should be done?"

High-level strategic objectives are observed in Business Analytics, so decisions should be strategy-based rather than purely data-driven. During the Kelsey-White Simulations for 4 years, decisions were made based on previous/history data. After seeing what happened, the model was optimized to a more sophisticated one. Prescriptive analytics has the potential to lead to automated decisions. Suppose someone has already coded what needs to be done in a specific case, and thus decisions are made automatically. This could result in dangerous extremes. As actions influence the world, new data is generated as a result. To demonstrate, a few students reduced the price of detergent for a year and then analyzed the results. It was noticed that the detergent began to sell more, and as an outcome, more transactions began to occur. As a result, more data was produced and needed to be analyzed.

Data, Data, Data

When making data-driven decisions in an organization, it is critical to understand how data is generated, collected, processed, stored, managed, visualized, and interpreted (Wing, 2019). Questions such as whether the source can be trusted, the value of the data, and if there was any interference or noise in the data. Should data be collected simply because one has the capability? Following the collection of data, are there any regulations or regulatory rules that may differ depending on the country or geographical region? Certain responsibilities may also accompany data collection, such as explicitly requesting personal data from individuals and then requesting permission to USE that data. When visiting any website in Ireland, for example, a pop-up window appears asking the user to enable cookies, whereas this is not the case when visiting the same website from India or the United States. As a result, it is clear how one aspect of data – data collection by an organization – affects society or the general population.

Sensors, GPS devices, mobile phones, social media, smart cars, roads, bridges, and buildings all generate data that can be analyzed and cross-referenced. Once the data has been collected, it must be saved. Once data has been stored, it must be managed. The hardware and software at each tier, as well as the tools used to perform the analytics, all require some level of management, just like DBMS. In turn, data management

introduces concepts such as redundancy, backup and restore, access rights, security, and validity (ACID properties), among others. Backup copies allow data to be restored from a previous point in time, assisting businesses in recovering from an unanticipated event. The health company was able to rebuild the systems after data loss, which was only possible because of data redundancy and backup, according to the lecture.

Various tools are used to uncover the potential of data during its journey to gain insights to support decision-making which is encapsulated with strategy. Entities with similar characteristics are identified and grouped together, and such attributes are measured in order to discover new entities that can be clustered for purposes, such as market research. Regression analysis is used to estimate the relationships between dependent and independent variables. Models are then fitted into the data. Using Big Data Analytics, analysts attempted to achieve accurate forecasts in strategic market analysis (Özemre and Kabadurmus, 2020). Thus, business analytics is a systematic way of thinking that employs qualitative, quantitative, and statistical computational tools and methods to analyze data, gain insights, inform, and support decision-making. (Power et al., 2018).

Incoming Data Scientists

Since the businesses are now dealing with information in unprecedented forms and volumes, there is a sudden appearance of data scientists in organizations where they make discoveries while swimming in data (Davenport and Patil, 2017). They assist decision-makers in transitioning from impromptu analysis to an ongoing dialogue with data. At times they may face technical difficulties, yet they present the insights in visually appealing formats to communicate the pattern.

It is critical to first understand how well the organization's process is and how analytics will help define that baseline, and then to define how analytics will help improve the process. After identifying the business problem, they must decide on what to focus on first, like whether any organization has investigated it before and, if so, whether a toolkit exists for it. It is critical to involve all stakeholders and determine what their needs are, how they will be impacted, and what challenges/difficulties they are facing. To see what is realistically possible and to understand what everyone's expectations are, as well as what is possible or not. In the very beginning, data science was only adopted in Adtech and MarTech industries. Facebook developed the Hive language for their Hadoop projects before releasing it to the open-source community. Many data scientists from data-driven businesses later contributed to and improved the toolkit. This summarizes Davenport's thinking, which consists of three steps: frame, solve, and act.

Saving the World with Nexleaf

“At Nexleaf, our mission is to partner with countries to ensure they have the data they need to build lasting solutions that improve the health of people.” (Nexleaf Analytics, n.d.)

Nexleaf is a not-for-profit organization which uses internet of things to save the world. They want to scale global health solutions with data. The co-founders, Nithya Ramanathan and Martin Lukac, wanted to use real-time data to help improve public health in every country, including the most remote areas of the world. Their initiatives are the Clean Cooking Program and the Vaccine Program. In the former, sensors generate data, which is analyzed and used to develop more effective and sustainable cookstove designs, fuel supply chains, and maintenance. Because most of their stakeholders are women, they do not expect them to adapt

to a new cooking solution, rather aim to develop such cooking solutions that meet the diverse needs of women. Nexleaf collaborates with partners to raise the bar for data because it is critical to have all aspects of data to derive accurate and precise insights (Nexleaf Analytics, n.d.). They have begun to face increasing competition from for-profit organizations, particularly in the Vaccine Program.

History and Importance of Data Visualization

Aside from analysis, data visualization aids in the presentation of results in a clear and simple manner that a human can easily understand and visualize. In 1786, William Playfair, abbreviated the modes of conveying information by inventing line, bar, area and pie charts (PLAYFAIR, 1798).

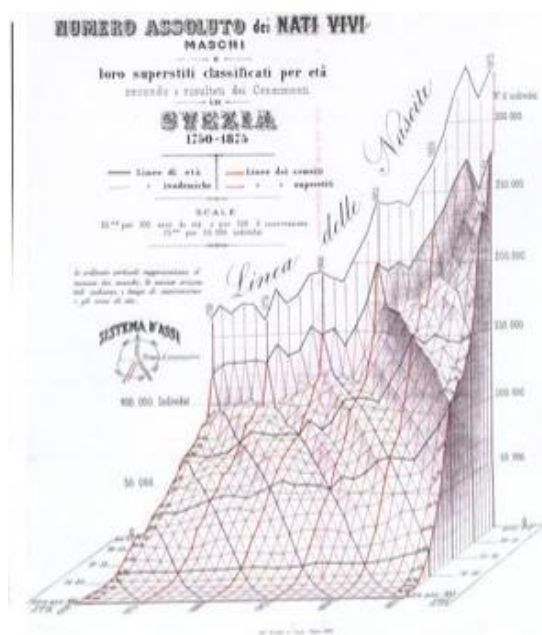


Fig 1: Luigi Perozzo – 3D visualization of the Swedish census

One of the first 3D representations (Fig-1) of data showing the age group of Swedish population between the 18th and 19th centuries, was produced by, Luigi Perozzo, an Italian statistician (Sridharan, 2017). Years are measured horizontally, numbers of individuals are measured vertically, and age groups (youngest nearest) are measured in depth as you move into and out of the image (Rau et al., 2017).

Numbers-based proof is essential, but numerical tables are never interesting. Florence Nightingale, the founder of modern nursing, understood this fact back in 1855 (www2.deloitte.com, n.d.). She was most notable for her presentation of the results. She organized care for wounded soldiers during the Crimean War. She started maintaining records of patients, wounds, diseases, and reasons of deaths. Nightingale recognized that soldiers were dying needlessly because of malnutrition, poor sanitation, and a lack of activity. She described graphs to show the difference in the number of deaths caused by diseases versus war wounds. In each area of the diagram, the color shading indicated the cause of death. The pinkish color represented deaths from battle wounds, the blueish-grey color represented deaths from infectious diseases, and the black color represented deaths from other causes. Post the shocking revelation that deaths were due

to unsanitary conditions rather than war, several reforms were introduced to improve the conditions. Later, she started campaigning for the same and became a celebrity.

From Visualization to Communication

The essence of analytical communication is describing the problem and its history, the model, the data used, and the relationships between the variables in the analysis. The findings should include the following: an outline of the research process, a summary of the findings and their implications, and a recommendation for action. It is usually best to start with the summary and recommendations, then communicate the findings in a meeting with the appropriate people or in a written, formal report.

Most researchers and professionals use scientific posters at conferences to visually present their findings. They present their work using mathematical and statistical terms. However, this will not function for the public because the audience is no longer quantitative or technical. As a result, these should be presented in terms of concepts that the audience will recognize and understand. In the case of for-profit organizations, the factor is usually the money that was earned or saved.

For example, in 2020, the entire world experienced unprecedented pandemic – Covid-19. Researchers, scientists, and organizations such as the World Health Organization used animations and visuals to emphasize the importance of masks, vaccinations, and isolation. They did not use technical terms such as Power-Law graphs – graphs with a significant number of vertices in higher degree – where the probability of a single connected component rises sharply as the average degree of the graph rises.

Little Bit Technical

According to the Power-Law graph, each person who has a disease (or information, etc.) passes it on to each of his contacts with an independent probability r . If the person has degree d , he can pass the disease to $d - 1$ contacts (excluding the one from whom he got the disease in the first place). As a result, the expected number of people infected by this person is $r(d - 1)$. R_0 (basic reproduction numbers) denotes the average number of contacts infected by a person based on a weighted average over vertices (Sean McGarraghy - MIS41160: Optimisation in Business). According to (Mousazadeh et al., 2021), all basic reproduction numbers (R_0) estimated in various studies were greater than 1, indicating that COVID-19 was still spreading in 2021. Of course, explaining the graphical model to society makes no sense. As a result, it is critical to communicate the essence through appealing visuals and simple language.

The researchers then presented frameworks based on Big Data Analytics to aid the government in its fight against epidemics. (Jia et al., 2020) suggests ways for governments to improve their epidemic response mechanisms, ranging from data collection to forecasting and reducing public fear by revealing the current epidemic situation and dispelling rumors.

Deceiving with Visuals

As important as it is to simplify the language for the general audience, it is primarily the analyst's responsibility to be aware of visual lies. The foundation is truth and evidence-based reporting. It is critical to prioritize information integrity over the sale of lies. The issue is that designers are either unaware or willingly disregard such principles. It is also possible that the designers are oblivious of technical aspects

such as basic statistics for detecting data distortion. To illustrate, lead author from (Lauer and O'Brien, 2020), taught data visualization class for 9 years to graduate and undergraduate students. She saw how students were willing to give up a lot of control to the software in order for it to perform the precise calculations to turn a spreadsheet of data into a visually appealing display. However, the defaults used by these programs would frequently adapt the display entirely to the data source. With axes being truncated and graph ranges are specific to the data being plotted, graphs from the same data set may produce similar-looking graphs with drastically different axis ranges and proportions. This made comparing data from different graphs misleading.

This poses a significant challenge for modern democracies because visual lies and/or misinterpretation may be confused with freedom of expression. There may be a lack of understanding in the fine line separating the two paradigms. Whether it is omitting datapoints that may highlight some trends that no longer fit the population or truncating the Y-axis, misleading the public to promote numbers is unethical. Such methods confuse rather than enlighten people. The issue isn't just a scarcity of genuine communicators. Therefore, reforms need to be adapted to potentially provide value and lead to a better organizational performance.

Need for Capability Framework

It is seen how technology, people and processes are involved in the data life cycle (data collection, processing, storage, analysis, interpretation) (Wing, 2019). The technologies revolve around data warehousing, visualization, interpretation tools, forecast modelling and so on (Watson 2010). From an organizational standpoint, various capabilities are required to involve and work on these because BA is a top priority for investment and defining BA capability framework becomes a prominent goal for organizations to address the gap in knowledge (Cosic, Shanks and Maynard, 2015). Davenport developed one of the first capability majority models (Davenport, 2014). As per Weill and Ross (2004), Lederner and Kayworth (2006), Davenport et al. (2007), and Negash (2004), 16 BA capabilities were identified and grouped into four capability areas based on their similarities: (i) governance, (ii) culture, (iii) people, and (iv) technology. This was done to identify high-quality BA capabilities that would increase value and competitive advantage. The thesis in Michael Lewis's book, *Moneyball*, of taking advantage of undervalued assets and selling overhyped slugging assets is a good example.

Incorporating ESR

While businesses are all about making money by increasing their profits and keeping ahead of their competitors, it is observed that possessing ethical behavior can increase productivity. Many people believe in facilitating societal change that benefits both businesses and society while also creating value (Pappas et al., 2018). Nexleaf Analytics is an appropriate example. Ethics does not require an additional layer and can and should be incorporated into all aspects of the data life cycle. However, we cannot rely solely on ethics, and thus regulation is required to protect against data breaches and hacking. Along with regulation comes the public's responsibility to read and comprehend the fine print (need to channel inner Amy Santiago from Brooklyn 99). On a more serious note, the EU General Data Protection Regulations established guidelines for organizations and individuals' privacy rights. Identifying the difference between objective and social responsibility is crucial (The Facebook Dilemma). The rules and regulations are nearly identical, but they may differ from one geographical region to another and from one organization to another. As AI grows in

power, more regulations are being put in place to protect humans, their safety, and their privacy, such as the EU AI Act Proposal (2021), leaving the birth of Ultron to a fictional Avengers world.

Organizations have begun to incorporate ethical decision making, such as regular trainings on what to do and what not to do, as well as how to handle situations in which ethics trumps leadership. With this holistic approach, organizational governance incorporates: (i) human rights, (ii) labor practices, (iii) environment, (iv) fair operating practices, (v) consumer issues and (vi) community involvement and development. For decision making, one should move away from "gut instinct" and towards a more empirical based approach. It is critical to be able to create value and competitive advantage while influencing decision making. Knowledge creation incorporates both judgment and experience.

Conclusion

Analytics is paving its way in the lives of the general audience. For the past week, most millennials, gen-z have been seen sharing their Spotify wrap for the year 2022, highlighting their top songs, artists, minutes spent listening to songs, and so on. With all the data generated and collected ethically and within regulations, it is reasonable to assume that people are interpreting the visuals and communicating the findings as a success. Of course, most of them are unaware of the work and capabilities of Spotify's data scientists and analysts. The need for and importance of data with layers of systems, processes, practices, standards, and finally wrapped in ESR can be seen for both business and society.

References

1. Caruso, J.V. (2018). Integrating Business Acumen and Analytics: A Simulation-based Approach. *Developments in Business Simulation and Experiential Learning: Proceedings of the Annual ABSEL conference*, [online] 45. Available at: <https://absel-ojs-ttu.tdl.org/absel/index.php/absel/article/view/3179> [Accessed 3 Dec. 2022].
2. Cosic, R., Shanks, G. and Maynard, S.B. (2015). A business analytics capability framework. *Australasian Journal of Information Systems*, [online] 19. doi:10.3127/ajis.v19i0.1150.
3. Davenport, T.H. and Harris, J.G. 2007. *Competing on Analytics: The New Science of Winning*. Harvard Business School Press, Boston.
4. Davenport, T. (2014). *Assessing Your Analytical and Big Data Capabilities*. [online] Available at: <https://www.tomdavenport.com/wp-content/uploads/2019/01/Assessing-Your-Analytical-and-Big-Data-Capabilities.pdf> [Accessed 3 Dec. 2022].
5. Davenport, T.H. and Patil, D.J. (2017). *Data Scientist: The Sexiest Job of the 21st Century*. [online] Harvard Business Review. Available at: <https://hbr.org/2012/10/data-scientist-the-sexiest-job-of-the-21st-century>.
6. Jia, Q., Guo, Y., Wang, G. and Barnes, S.J. (2020). Big Data Analytics in the Fight against Major Public Health Incidents (Including COVID-19): A Conceptual Framework. *International Journal of Environmental Research and Public Health*, [online] 17(17), p.6161. doi:10.3390/ijerph17176161.
7. Lauer, C. and O'Brien, S. (2020). The Deceptive Potential of Common Design Tactics Used in Data Visualizations. *Proceedings of the 38th ACM International Conference on Design of Communication*. doi:10.1145/3380851.3416762.
8. Leidner, D. E., and Kayworth, T. 2006. A review of culture in information systems research: Toward a theory of information technology cultural conflict, *MIS Quarterly*, 30:2, 357-399.
9. Mousazadeh, M., Naghdali, Z., Rahimian, N., Hashemi, M., Paital, B., Al-Qodah, Z., Mukhtar, A., Karri, R.R., Mahmoud, A.E.D., Sillanpää, M., Dehghani, M.H. and Emamjomeh, M.M. (2021). *Chapter 9 - Management of environmental health to prevent an outbreak of COVID-19: a review*. [online] ScienceDirect. Available at: <https://www.sciencedirect.com/science/article/pii/B978032385780200007X> [Accessed 4 Dec. 2022].
10. Negash, S. 2004. Business Intelligence,” *Communications of the Association for Information Systems*, 13, 177-195.

11. Özemre, M. and Kabadurmus, O. (2020). A big data analytics based methodology for strategic decision making. *Journal of Enterprise Information Management*, ahead-of-print(ahead-of-print). doi:10.1108/jeim-08-2019-0222.
12. Pappas, I.O., Mikalef, P., Giannakos, M.N., Krogstie, J. and Lekakos, G. (2018). Big data and business analytics ecosystems: paving the way towards digital transformation and sustainable societies. *Information Systems and e-Business Management*, 16(3), pp.479–491. doi:10.1007/s10257-018-0377-z.
13. PLAYFAIR, W. (1798). *Lineal Arithmetic; applied to shew the progress of the commerce and revenue of England during the present century, which is represented ... by thirty-three copperplate charts, etc.* [online] Google Books. Available at: <https://books.google.ie/books?id=7fBZAAAACAAJ&pg=PA7&lpg=PA7&dq=%E2%80%9Cthe+best+judge+of+proportion> [Accessed 4 Dec. 2022].
14. Power, D.J., Heavin, C., McDermott, J. and Daly, M. (2018). Defining business analytics: an empirical approach. *Journal of Business Analytics*, [online] 1(1), pp.40–53. doi:10.1080/2573234x.2018.1507605.
15. Sridharan, M. (2017). *Data Visualization - History And Origins*. [online] thinkinsights.net. Available at: <https://thinkinsights.net/data-literacy/data-visualization-history/#John-Snow-Map-of-Cholera-outbreaks> [Accessed 4 Dec. 2022].
16. Watson, H.J. 2010. BI-based Organizations. *Business Intelligence Journal*, 15(2), 4.
17. Weill, P. and Ross, J.W. 2004. *IT Governance: How Top Performers Manage IT Decision Rights for Superior Results*. Harvard Business School Publishing, Boston, MA.
18. Wing, J.M. (2019). The Data Life Cycle. *Issue 1*. doi:10.1162/99608f92.e26845b4.
19. www2.deloitte.com. (n.d.). *Telling a story with data | Deloitte Insights*. [online] Available at: <https://www2.deloitte.com/us/en/insights/deloitte-review/issue-12/telling-a-story-with-data.html>.

Footnotes

1. For fig1, the picture is extracted from the book: Rau, R., Bohk-Ewald, C., Muszyńska, M.M. and Vaupel, J.W. (2017). *Visualizing Mortality Dynamics in the Lexis Diagram*. [online] Google Books. Springer. Available on page 23 of: <https://books.google.ie/books?hl=en&lr=&id=ttpCDwAAQBAJ&oi=fnd&pg=PR5&dq=Luigi+P+erizzo> [Accessed 4 Dec. 2022].