Option #1: Business Analytics Program

Scott Miner

Colorado State University – Global Campus

**Option #1: Business Analytics Program**

For his *Portfolio Project* for *MIS542 – Business Analytics,* the student will evaluate the business analytics program at an institution in the education industry, Institution A. Following the trend in education to move more courses online, Institution A is amassing ever-increasing amounts of learning-related data. Institution A wants to extract value from these big data sets and optimize online learning opportunities (Ferguson, 2012). As the cornerstone of his analysis, the student uses Figure 11.1 and Table 11.1, as presented by Nelson (2018), which emphasize the importance of balancing effectiveness and efficiency in a business analytics program. *Effectiveness* refers to doing the right thing to create value, whereas *efficiency* refers to generating maximal productivity with the least possible expense. The student divides his analysis of the business analytics program into the five following stages: (a) *problem framing*, (b) *data sensemaking*, (c) *analytics model development*, (d) *results activation*, and (e) *analytics product management* (Nelson, 2018). Lastly, within each of these five sections, the student includes a plan of action to help Institution A improve its business analytics program.

**Problem Framing**

Often, analysts within the Institutional Research (IR) department at Institution A skip the problem framing step and jump right into the data collection phase, overly eager to solve a problem before adequately defining it. In the *problem framing* stage of Table 11.1, Nelson (2018) outlines several ways for organizations to increase the effectiveness of their analytics programs, one of which includes prioritizing the right problem to solve. To prioritize the problem, Institution A should follow the five processes of problem framing as outlined by Nelson: (a) *problem definition,* (b) *root cause investigation,* (c) *hypothesis generation,* (d) *question design,* and (e) *business case prioritization.* In essence, Institution A needs to define the problem, whom it affects, its impact, and what a successful solution looks like, all of which comprise the **problem statement.** In defining the problem, analysts at Institution A might state that the institutionis experiencing a shortfall in projected student retention rates*.* After all, Shapiro *et al.* (2016, as cited in Larrabee Sønderlund *et al*., 2019) describe that nearly 8 million (39%) of the 20.5 million US students enrolled in university courses will drop out before graduating. Likewise, the College Board Advocacy & Policy Center (2010, as cited in Jayaprakash *et al.,* 2014) asserts that the US ranks 12th globally in the percentage of 25- to 34-year-olds with an associate’s degree or higher. In effect, there is tremendous potential for Institution A to increase student retention rates.

To solve this problem, analysts at Institution A should subject it to *a root cause investigation* to determine why it exists. Strategies in this area outlined by Nelson (2018) include using affinity and fishbone diagrams to conceptualize problem causes by type, also known as **category brainstorming.** For instance, Frankola (2002, as cited in West *et al*., 2015) describes potential causes of student attrition, including difficulties with technology and social isolation. Likewise, Mazza and Dimitrova (2004) describe how students may become disoriented in course hyperspaces and lose motivation. Lastly, Cambruzzi *et al.* (2015) describe how students may fail to see the importance of course material. All these examples represent potential causes of why students may drop out of their studies.

Nelson (2018) presents the next step within problem framing as *hypothesis* *generation*, where we utilize **divergent thinking** to develop ideas about how we believe the world works. We work in teams in this step, employing sticky notes to generate hypotheses that allow us to translate our belief statements into SMART goals. For instance, in the case of Institution A, we might say that by implementing learning analytics (LA) for a student, Andrew, we will increase student retention rates, which we will know to be true when they have increased by 20%. Likewise, Krumm et al. (2014) found that courses implementing a LA software called *Course Signals* (CS), which uses multiple dynamic indicators to convey academic performance information to students, increased retention rates by approximately 10 percent to 25 percent compared to courses not using the program. We utilize a 2 x 2 matrix to evaluate the risk and impact of our hypothesis (Nelson, 2018).

We formulate the problem into a question and generate potential study designs to answer our question during the *Question Design* phase. For instance, we could use simple regression models to examine associations between student demographics and retention rates. Milliron et al. (2014) implemented an example of A/B testing, finding that students at-risk of failure who received a phone or email intervention were 3.21%—7.62% more likely to remain enrolled than students in the control group. Nelson (2018) recommends having a control group when possible. Notably, Milliron et al. also found that phone calls were more effective for early-term students, while email was more effective for students with 10+ terms at the institution. In the last phase of problem framing, *Business Case Prioritisation,* Nelson suggests a project prioritization matrix to assess the importance of solving a problem. In this step, we create the **business case**,whichis a justification for the proposed project according to its expected benefits and describes what the project does and how it impacts the organization’s strategic business objectives.

**Data Sensemaking**

Referring back to Table 11.1 of the text, Institution A’s biggest challenge in the data sensemaking stage relates to efficiency and utilizing the most appropriate technology to eliminate waste and avoid redundancy (Nelson, 2018). Nelson (2018) divides this best practice area into four processes (a) *data identification and prioritization*, (b) *data collection and preparation*, (c) *data profiling and characterization*, and (d) *visual exploration*. In *data identification and prioritization,* Nelson suggests mocking up the data if it is not readily available to aid in testing its utility. Institution A can access its data warehouse via web-based tools but could benefit from updated technology, as it relies primarily on Excel. For instance, three case studies conducted by Milliron et al. (2014) utilized both Civitas Learning’s Illume software, which is a predictive analytics platform, and its Inspire application, which is an action analytics application that uses design thinking to leverage insight analytics in easy to consume formats for administrators. The utilization of both these applications allowed Milliron et al. to conduct their case studies in three different universities more efficiently. Updated software would also help in the *data collection and preparation* phase, where we use **data integration** tools to extract, transform, and process data for analytics. Nelson also suggests implementing a **sandbox** environment to help with computationally expensive transformations. **Source-to-target mapping** will aid in documenting the necessary data transformations.

New technologies could also aid Institution A in the *data profiling and characterization* phase, where we explore the data and its patterns, examining the technical and business quality. *Technical quality* refers to concepts like invalid or missing data, whereas *business quality* encompasses concepts like relevance and completeness. When we characterize the data, if seeking to understand a relationship between two or more variables, we classify them into the **response** (e.g., outcome) and **explanatory** (e.g., predictor)variables. For instance, the Student Probability Model (SPM) implemented by the Open University UK (OU) predicts whether students will reach specific milestones based on explanatory factors, including the number of fails and total credits (Herodotou et al., 2019). Lastly, in *visual exploration,* we use **visual data discovery** to find patterns in the data set.

**Analytics Model Development**

In *Analytics Model Development,* Nelson (2018) covers a broad range of topics, dividing this best practice area into four processes: (a) *making comparisons,* (b) *measuring associations,* (c) *making predictions,* (d) *detecting patterns.* When referring to Table 11.1, one-way Institution A could improve its effectiveness in this area is by testing hypotheses for their validity. Nelson describes making comparisons as one of the most common activities in analytics. In this area, we may use statistical tests such as the F-test or T-test to compare differences between means or variances. However, when doing so, it is important to state our null and alternative hypotheses. Then, we check if our data provide evidence against the null hypothesis. For instance, Milliron et al. (2014) found that 3% more students completed their courses when receiving LA interventions compared to those who did not. The results were statistically significant.

The same concepts apply in the process area of *measuring associations*. For instance, Fritz (2011) hypothesized that students more active in the course management system (CMS) at the University of Baltimore, Maryland’s would receive higher grades. Results indicated a positive correlation between grades and online activity. Students who earned a D or F used the CMS 39% less than students receiving a C grade or higher. *Making predictions* refers to developing advanced analytical models to predict the value of one variable given another. At Institution A, we may wish to predict at-risk students or letter grades. The SPM implemented at OU uses logistic regression to predict whether students reach specific milestones, a categorical response variable (Herodotou et al., 2019). Finally, *pattern detection* is nearly synonymous with machine learning, assigning labels to objects. Lykourentzou et al. (2009) used machine learning techniques to classify students likely to drop out of e-learning courses and achieved significantly better results than those reported in the relevant literature. As such, Institution A may wish to apply any of these examples to its business analytics program to efficiently and effectively identify students at risk of dropping out.

**Results Activation**

To improve effectiveness in the *results activation* area, Nelson suggests mastering storytelling for influence, and to improve efficiency,he suggests documenting our findings and knowledge for reuse. Institution A could benefit from both of these recommendations. Nelson divides this stage into three processes: (a) *solution evaluation*, (b) *operationalization*, and (c) *presentation and storytelling*. In the solution evaluation process, we review and validate our model, evaluate our results, and assess our model’s impact. The area under the receiver operating characteristic (ROC) curve is a useful visualization that helps convey model robustness. Milliron *et al.* (2014) used this visualization to explain how their studies model provided a 141% improvement in accurately identifying eventual non-persistent students for intervention compared to randomly reaching out to students. As we assess our model’s impact, we want to uncover potential process, technology, organizational, and people impacts.

Then, we can operationalize our analytics model by determining the most appropriate strategy for deployment. Herodotou *et al.* (2019) provide several recommendations for the adoption of predictive learning analytics (PLA) in HE, including providing evidence of the effectiveness of PLA, encouraging communication across stakeholders, setting a shared PLA vision, considering ways to mitigate teacher resistance, and using PLA to complement, not replace, the teaching practice. Likewise, Larrabee Sønderlund et al. (2019) recommend breaking all outcomes down by relevant population characteristics, recording student and faculty experiences of the initiatives, and standardizing the assessment of intervention programs. Nelson (2018) confirms these recommendations, adding that we need to define key metrics and evaluate our programs. For Institution A to effectively and efficiently master storytelling, Nelson suggests using visualizations to bring a story to life and use a story blueprint to understand what we want to convey. From that point on, we tell our data stories, intending to create a sense of urgency and a clear pathway to action.

**Analytics Project Management**

Finally, to improve Institution A’s effectiveness within the *analytics project management* area, Nelson (2018) suggests defining the program evaluation strategy for analytics products and, to increase efficiency, suggests capturing knowledge for future use. Nelson divides this best practice into five process areas: (a) *value management,* (b) *analytics lifecycle execution,* (c) *quality processes,* (d) *stakeholder engagement and feedback¸* (e) *capability and talent development.* Within the *value management* process, analytic product managers want to maintain alignment with organizational strategy, lead the product management culture, define the quality of analytics products, and promote value. Nelson emphasizes the importance of maintaining a highly transparent process regarding how we handle analytic requests, rather than adhering to strict prioritization guidelines, to engender trust between analytics teams and customers.

Additionally, product managers should define what it means to deliver a quality analytics product. Endel and Piringer (2015) state that researchers often neglect documentation in scientific research. However, in defining quality data products, documentation is essential. Within the *analytics lifecycle execution* area, product managers are responsible for the analytics processes that produce value. **Product lifecycle management** encompasses six vital steps: (a) opportunity identification, (b) design, (c) testing, (d) launch, (e) assessment, and (f) retirement. Institution A currently utilizes a single “decider” who triages product requests. Nelson advises against this approach and recommends an analytics product prioritization matrix to promote collaboration. Essentially we want to **right-size** analytics product requirements appropriately. In the design phase, we want to *prototype* the minimum viable product (MVP) to address the customers’ needs before testing. After launching a solution, we should assess product performance. Likewise, Larrabee Sønderlund *et al.* (2019) suggest developing standardized assessments of LA intervention programs and recording student and faculty experiences of the initiatives. While neglected in Institution A’s current business analytics program, retirement planning is essential because analytics products have an expected lifespan. As mentioned previously, during *quality processes,* product managers should document all analytics product processes. The **analytics quality** of the data productassesses the product’s overall abilityto serve its function. We want to engage stakeholders early and often and document changes to analytics products. Herodotou *et al.* (2019) emphasize the importance of bidirectional communication by recommending that organizations allocate resources for communication across stakeholders. Analytics done well leads to change, and Herodotou *et al.* also recommend considering ways to mitigate teachers’ resistance to change when implementing PLA initiatives. Lastly, in the *capability and talent development* process, we want to create a supportive learning culture. Moreover, if we understand where our people are regarding analytics competencies and the organization’s strategic goals, we can use this framework to integrate professional development plans.

**Conclusion**

References

Cambruzzi, W. L., Rigo, S., & Barbosa, J. L. V. (2015). Dropout Prediction and Reduction in Distance Education Courses with the Learning Analytics Multitrail Approach. *J. Univers. Comput. Sci.*

Endel, F., & Piringer, H. (2015). Data Wrangling: Making data useful again. *IFAC-PapersOnLine*, *48*(1), 111–112. <https://doi.org/10.1016/j.ifacol.2015.05.197>

Ferguson, R. (2012). Learning analytics: Drivers, developments, and challenges. *International Journal of Technology Enhanced Learning*, 4(5/6), 304. <https://doi.org/10.1504/IJTEL.2012.051816>

Fritz, J. (2011). Classroom walls that talk: Using online course activity data of successful students to raise self-awareness of underperforming peers. *The Internet and Higher Education*, *14*(2), 89–97. <https://doi.org/10.1016/j.iheduc.2010.07.007>

Herodotou, C., Rienties, B., Verdin, B., & Boroowa, A. (2019). Predictive Learning Analytics “At Scale”: Guidelines to Successful Implementation in Higher Education. *Journal of Learning Analytics,* 6(1). https://doi.org/10.18608/jla.2019.61.5

Jayaprakash, S. M., Moody, E. W., Lauría, E. J. M., Regan, J. R., & Baron, J. D. (2014). *Early Alert of Academically At-Risk Students: An Open Source Analytics Initiative. Journal of Learning Analytics,* 1(1), 6–47.

Krumm, A. E., Waddington, R. J., Teasley, S. D., & Lonn, S. (2014). A learning management system-based early warning system for academic advising in undergraduate engineering. In *Learning analytics* (pp. 103–119). Springer.

Larrabee Sønderlund, A., Hughes, E., & Smith, J. (2019). The efficacy of learning analytics interventions in higher education: A systematic review. *British Journal of Educational Technology*, 50(5), 2594–2618. <https://doi.org/10.1111/bjet.12720>

Lykourentzou, I., Giannoukos, I., Nikolopoulos, V., Mpardis, G., & Loumos, V. (2009). Dropout prediction in e-learning courses through the combination of machine learning techniques. *Computers & Education*, *53*(3), 950–965. <https://doi.org/10.1016/j.compedu.2009.05.010>

Mazza, R., & Dimitrova, V. (2004). Visualising Student Tracking Data to Support Instructors in Web-Based Distance Education. *13th International World Wide Web Conference (WWW 2004) - Educational Track. 2004 May 17-22*, 154–161.

Nelson, G. S. (2018). *The analytics lifecycle toolkit: A practical guide for an effective analytics capability*. Wiley-Blackwell.

West, D., Heath, D., & Huijser, H. (2015). Let’s Talk Learning Analytics: A Framework for Implementation in Relation to Student Retention. *Online Learning*, 20(2). <https://doi.org/10.24059/olj.v20i2.792>

Zilvinskis, J., & Wills, III, J. (2019). Learning Analytics in Higher Education: A Reflection. *InSight: A Journal of Scholarly Teaching*, 14, 43–54. <https://doi.org/10.46504/14201903zi>