University Opportunities, Abilities and Motivations to Create Data Analytics Programs

Troy J. Strader
Drake University, Troy.Strader@drake.edu

Andrew Bryant
Drake University, Andrew.Bryant@Drake.edu

Follow this and additional works at: http://aisel.aisnet.org/jmwais

Recommended Citation
Strader, Troy J. and Bryant, Andrew ( ) "University Opportunities, Abilities and Motivations to Create Data Analytics Programs," Journal of the Midwest Association for Information Systems (JMWAIS): Vol. 2018 : Iss. 1 , Article 4.
DOI: 10.17705/3jmwa.000039
Available at: http://aisel.aisnet.org/jmwais/vol2018/iss1/4
University Opportunities, Abilities and Motivations to Create Data Analytics Programs

Troy J. Strader
Drake University, troy.strader@drake.edu

Andrew Bryant
Drake University, andrew.bryant@drake.edu

Abstract

Some US colleges and universities have developed undergraduate and graduate data analytics programs in the past five years, but not all universities appear to have sufficient resources and incentives to venture into this multidisciplinary academic area. The purpose of this study is to identify the characteristics of schools that have developed data analytics programs. The study utilizes the motivation-ability-opportunity (MAO) theoretical framework to identify factors that increase the likelihood that a university will develop a data analytics program. An analysis of 391 regional master’s universities in the US finds that schools with data analytics programs are more likely to be in larger cities and have larger student enrollments, better educational quality rankings, and an existing statistics and/or actuarial science program. These findings support the idea that data analytics programs are more likely to be created when universities have opportunities to access a larger number of businesses and governmental organizations, and sufficient resources to support program development, while also having abilities associated with innovation and faculty resources. Preliminary results also indicate that there are two motivations – need to increase student enrollment and need to maintain an up-to-date curriculum.

Keywords: Data analytics education, business analytics, data science, motivation-ability-opportunity

DOI: 10.17705/3jmwa.000039

Copyright © 2018 by Troy J. Strader and Andrew Bryant
1. Introduction

It is estimated that by 2020 about 1.7 megabytes of data will be created each second for every person on earth (Marr, 2015). This data includes pictures and videos to document one’s life, the details of every transaction across every global organization’s supply chain, and every interaction a government agency has with its citizens. Businesses and government agencies are increasingly viewing their data as an asset that can be utilized to better serve their customers and improve their competitive position, but identifying value in vast amounts of data requires skills that may not be readily available in their current employees. This situation has led to increased demand for data analysts across a wide range of industries and governmental functions. It is difficult to forecast future demand for data analysts, but one prediction is that there will be a 50-60% shortfall for supply of data analysts in the US by 2018 (Orihuela & Bass, 2015). This scenario provides opportunities for current and future students as well as educational institutions at all levels.

Growth in data analytics jobs has led to an increase in student interest for analytics degree programs. The natural response has been an increase in development and implementation of undergraduate and graduate analytics programs in both small and large US colleges and universities (Chiang, Goes, & Stohr, 2012). However, analytics programs can only be found in a relatively small number (approximately 9%) of colleges and universities in the US (Tableau, 2016). Most of these academic programs are relatively new, but there is sufficient activity to reflect on how the development of data analytics programs has evolved. The purpose of this study is to address the following questions:

1. What external and internal opportunities increase the likelihood that a university will create a data analytics related program?
2. What internal abilities increase the likelihood that a university will create a data analytics related program?
3. Given a university’s opportunities and abilities, what motivates them to develop a data analytics related program?

This study addresses issues in two research areas – the current state of business intelligence and analytics education, and the design and development of business intelligence and analytics curriculum (Wang, 2015). The following sections describe the study’s background, hypotheses and findings. First, the motivation-ability-opportunity theoretical framework is discussed and hypotheses are identified. Next, the methodology and data are described followed by a presentation of the study’s findings. Finally, the paper concludes with a discussion of overall conclusions, implications, and directions for future research.

2. Background and Study Hypotheses

The motivation-ability-opportunity (MAO) theory has been utilized in a variety of studies related to understanding consumer attitudes, information processing and decision-making. Early studies looked at the impact of consumer motivations, abilities, and opportunities on advertising effectiveness as part of the Elaboration Likelihood Model (Batra & Ray, 1986; Cacioppo & Petty, 1984; MacInnis, Moorman, & Jaworski, 1991). This theory states that motivation, ability, and opportunity are unique concepts that are necessary conditions for effortful processing of information (Andrews, 1988). More recently, the theory has been used as a theoretical framework for categorizing antecedent variables in studies of a wide variety of decision-making contexts. The following provides some examples of studies in this area.

The MAO framework was used to explain knowledge sharing among employees (Siemsen, Roth, & Balasubramanian, 2008). The authors developed a constraining factor model that identified MAO-related bottlenecks that determined the degree of knowledge sharing. In another study, the MAO framework was used to develop and test a model to identify the driving factors behind airline ticket purchases (Bigné et al., 2010). And in another study, it was found that motivations, abilities, and opportunities influenced physician’s adoption of electronic medical records (Govindaraju, Hadining, & Chandra, 2013). MAO has also been proposed as a framework for organizing the factors that impact electronic market participation by looking at the interrelationship between consumer opportunities to access the electronic market, their abilities to participate in this new channel, and finally, their motivations to purchase online rather than through traditional channels (Strader & Hendrickson, 1999).

Given the MAO framework’s usefulness for organizing and categorizing issues addressed in a study, it is incorporated into this study as the basis for identifying sets of factors that may influence university decisions to develop and implement data analytics related academic programs. In this study we adapt the three components of the MAO framework, but they are addressed by first considering external and internal opportunities and then looking at internal abilities. Internal
motivations are addressed separately in the second stage of this study. The following provides a summary of the study’s issues and related hypotheses.

2.1 External and Internal Opportunities

The first issue is whether a university has the external and/or internal resources to even consider developing a data analytics program. Externally, larger cities will have more businesses and government agencies to provide support for data analytics educational programs and jobs. Some examples would include opportunities for program funding, data, and funding for real-world projects, and opportunities for student internships and full-time employment. Universities in small communities would not have as much access to these resources. As a result, the physical location of a university could enhance the viability of a data analytics program thus increasing its appeal to the university stakeholders. This leads to the first hypothesis.

H1. Universities in larger cities are more likely to have data analytics programs.

The next issue considered is the internal resources that universities have to support the development of data analytics programs. Universities with larger student enrollments are far more likely to have financial and personnel resources to develop new programs when compared with very small universities. In other words, launching a new program at a smaller university represents a bigger commitment in terms of resources when compared with a larger university. As such, it would be less risky for a larger university to launch a new program like data analytics. This is described in hypothesis 2.

H2. Universities with larger enrollments are more likely to have data analytics programs.

2.2 Internal Abilities

In addition to variation in the opportunity to develop a new program, universities will vary based on their internal abilities to convert existing assets into developing an appropriate and innovative academic program. Schools without these existing assets may have too many hurdles to overcome to consider developing a data analytics program even if the opportunity presents itself.

The first university-related ability relates to their overall academic quality. A large component of university rankings are related to resources, including resource availability to faculty (USnews.com, 2017). As a result, universities with better academic rankings (i.e. lower ranking numbers) can be more innovative than universities that have lower overall academic quality. This relative advantage, in terms of ability, increases the likelihood of the development of new curricula to keep up with changing market needs and student demand. Therefore, H3 is described below.

H3. Universities with better overall academic rankings are more likely to have data analytics programs.

The final two hypotheses consider the existing assets a university may have available from existing majors in areas associated with data analytics, specifically statistics and/or computer information systems. If schools already have faculty and degree programs in these areas, then there is a shorter path to developing new data analytics programs. If a school does not have either of the programs then it would be much more difficult to start a new program given the need for funding, administrative support, and hiring new faculty. Most universities have math and computer science programs so these related areas would not be differentiators when comparing universities. This leads to the final two hypotheses.

H4. Universities with existing statistics or actuarial science majors are more likely to have data analytics programs.

H5. Universities with existing information systems majors are more likely to have data analytics programs.

The study’s opportunity and ability factors and related hypotheses are summarized in Figure 1.
3. Study Methodology

To address the opportunity and ability factors described above, a sample of US colleges and universities was identified. The sample includes all of the regional master’s universities identified in the US News rankings for 2016 that were ranked 100 or better in each of four regions (Usnews.com Midwest, 2016; Usnews.com North, 2016, Usnews.com South, 2016; Usnews.com West, 2016). Previous research into offerings of analytics programs found the majority of the national universities identified by US News offered some form of analytics program (Tableau, 2016). However, these universities tend to be large research oriented institutions. Universities in the regional master’s category offer a full range of undergraduate degree programs and some master’s degrees. These universities do not offer many doctoral-level programs. This sample provides a diverse set of small and large universities located in a variety of different sized communities which allows for the testing of the hypotheses previously listed.

To study H1-H5, data was collected from three different sources: The US News rankings web pages, each university’s web page, and the US census. The US News ranking web pages provided information about each universities region, ranking, and undergraduate enrollment. Next, each university website was reviewed to identify which universities had a data analytics related undergraduate or graduate degree program. This data was typically found through links on an academics page or through a search using the terms data analytics, business analytics, or data science. In addition, it was also determined whether each university had existing majors in actuarial science, statistics, or information systems. Finally, population data was identified for each university’s city (or metropolitan area) from a US census website (Census.gov, 2013). The combination of data collected for each ranked university was compiled and analyzed in SAS v9.4. The sample characteristics are summarized in Table 1.

A descriptive analysis of the data collected for each university showed that 56 (14%) of the universities have an undergraduate data analytics related major. The most common major names are data science, business analytics, and bioinformatics, but there are several other names used to describe these programs. A review of the websites also showed that there are 58 (15%) schools with graduate (master’s level) programs. Again, some of the most common program names...
are data science or business analytics, but there are also a number of bioinformatics programs. Interestingly, only 9 (less than 3%) of the 391 universities reviewed have both an undergraduate and graduate data analytics program. It appears that the most common strategy is to focus on either the undergraduate or graduate program, but not both.

4. Findings

A logistics regression model was used to assess H1 through H5. The logistic regression model was used after a comparison against a decision tree. The logistic regression model was found to have a superior fit using the area under a ROC curve (Logistic Regression = 0.767 vs. Decision Tree = 0.720). The independent variable parameter estimates are shown in Table 2. In addition, as universities are ranked by region (e.g. each region can have a top ranked university), dummy variables were included for each region and the West region was treated as the baseline for comparison. The modeled dependent variable is a binary variable where the value of one indicates that a school has either an undergraduate major in data analytics, or a graduate level master’s degree program in data analytics, or both. Zero indicates that a university does not have a data analytics program at any educational level. In other words, we are modeling the likelihood a university will have a data analytics program in some form versus a university will not have any data analytics program.

In total, the variables in our model enhance the overall model fit. A likelihood ratio test of the full model against an intercept only model indicates that, as a set, these variables significantly increase the model fit ($\chi^2 = 70.21$, d.f. = 8, $p < 0.001$). Further, a Receiver Operator Characteristic (ROC) curve, shown in Figure 2, was used to examine the model fit by plotting the true positive classification (sensitivity) against false positives (1 - specificity). The area under our ROC curve was 0.77 indicating a good model fit. For comparison sake, the area under a perfect model ROC curve equals one. Finally, Tjur’s coefficient of discrimination (2009) was calculated (d = 0.18). This statistic indicates a benefit associated with our model when comparing predicted probabilities between universities with and without data analytics programs.
In order to decrease the number of leading zeros and ease the reporting of our model’s parameter estimate, undergraduate enrollments were scaled by 1,000 and city population was scaled by a 1,000,000. The model shows that universities with larger city populations ($\beta = 0.08, \text{Wald } \chi^2 = 8.06, \text{d.f.} = 1, p < 0.01$) and larger undergraduate enrollments ($\beta = 0.04, \text{Wald } \chi^2 = 4.03, \text{d.f.} = 1, p < 0.05$) are significantly more likely to have a data analytics program. This supports H1 and H2 and indicates that greater opportunity, both internal and external, is positively related to the presence of data analytics.
The model also shows that there is a statistically significant negative relationship ($\beta = -0.02$, Wald $\chi^2 = 22.24$, d.f. = 1, $p < 0.01$) between university rank and the existence of a data analytics programs at a university. Universities with better overall academic quality (lower ranking numbers) are more likely to have some type of data analytics program (either undergraduate, graduate, or both). As a decrease in rank number indicates a better overall ranking, this result provides support for H3. Finally, it is found that universities with statistics and/or actuarial science majors are significantly ($\beta = 0.92$, Wald $\chi^2 = 9.79$, d.f. = 1, $p < 0.01$) more likely to have data analytics programs. This supports H4. The significance of H3 and H4, taken together, indicate potential pathways of how a university’s ability is related to the likelihood a data analytics program will be implemented.

Interestingly, the relationship between information systems majors and data analytics programs is positive but not statistically significant ($\beta = 0.51$, Wald $\chi^2 = 2.74$, d.f. = 1, $p = 0.098$). H5 is not supported. One explanation could be that some information systems programs have evolved into an analytics related program. Information systems was available prior to the development of the data analytics program, but may no longer be offered as a separate major program. Thus, information systems could provide resources to support the development of a data analytics program, but the data only provides a snapshot of the situation at this one point in time and cannot capture whether information systems programs provided any essential resources behind the development of the data analytics program.

Our model shows support for H1-H4, while H5 is directionally consistent but it is not statistically significant. As a whole the results show that universities in larger cities, with larger enrollments and better overall academic rankings, and existing majors in related statistics and quantitative areas, are more likely to have a data analytics program, ceteris paribus.

In order to further examine the robustness of our findings, we test our findings under alternate specifications. Specifically, we begin with an intercept only model then sequentially add variables to our model. The intention is to show how consistent the parameter estimates are after controlling for other variables. These results can be seen in Table 3. Also included in Table 3 is each model’s Akaike Information Criterion (AIC), a value used to compare models such that a smaller number indicates a better fit.

With the exception of the variable undergraduate enrollment, parameter estimates appear consistent under alternate model specifications. The addition of other variables to the model does not appear to have much effect on the overall conclusion to the hypotheses tests in that the signs of the parameter estimates are reliable across the different models. The consistency of the results in both direction and magnitude indicate a robust model and add to the strength of our conclusions.

5. Internal Motivations

An additional issue addressed in this study is to dig deeper into the reasons why some universities have developed undergraduate data analytics majors. To answer this question a short exploratory questionnaire was sent to faculty responsible for administering undergraduate data analytics majors. The focus was on the more common data science and business analytics majors. Questionnaires were not sent to administrators for bioinformatics programs. Questionnaires were sent to 35 program administrators and eight responses (23% response rate) were received. Each administrator was asked to identify what motivated their university to develop their undergraduate data analytics program. Results are summarized in Table 4.

Qualitatively examining the eight responses it appears that there are two primary motivators for the development of these programs. The primary reasons stated for why programs were created was that there was a need to increase undergraduate student enrollment and a need to maintain an up-to-date curriculum. Universities were motivated to meet these needs because it enhanced their tuition revenue and maintained or enhanced their program offerings and overall academic quality.
Table 3. Parameter Estimates to Test Hypotheses

<table>
<thead>
<tr>
<th>Parameter (modeled DV = 1)</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
<th>Model 5</th>
<th>Model 6</th>
<th>Full Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-1.10</td>
<td>-1.37</td>
<td>-1.65</td>
<td>-0.68</td>
<td>-0.92</td>
<td>-1.22</td>
<td>-1.44</td>
</tr>
<tr>
<td>City Population (in millions)</td>
<td>0.10</td>
<td>0.09</td>
<td>0.09</td>
<td>0.09</td>
<td>0.09</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td>Undergraduate Enrollment (in thousands)</td>
<td>0.05 **</td>
<td>0.06</td>
<td>0.04</td>
<td>0.03</td>
<td>0.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>University Ranking</td>
<td>-0.02</td>
<td>-0.02</td>
<td>-0.02</td>
<td>-0.02</td>
<td>-0.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Statistics and/or Actuarial Science Major</td>
<td>0.99 **</td>
<td>0.97</td>
<td>0.92</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Information Systems Major</td>
<td></td>
<td></td>
<td></td>
<td>0.50</td>
<td>0.51</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Region: North (vs. West)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.74</td>
<td></td>
</tr>
<tr>
<td>Region: South (vs. West)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td>Region: Midwest (vs. West)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.12</td>
<td></td>
</tr>
<tr>
<td>AIC</td>
<td>442.30</td>
<td>427.38</td>
<td>423.05</td>
<td>397.82</td>
<td>387.91</td>
<td>387.09</td>
<td>388.08</td>
</tr>
</tbody>
</table>

***p < .01; **p < .05; *p < .10

Table 4. Motivations to Develop an Undergraduate Data Analytics Program

<table>
<thead>
<tr>
<th>Motivations</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Need to increase undergraduate student enrollment</td>
<td>4</td>
</tr>
<tr>
<td>Need to maintain an up-to-date curriculum</td>
<td>6</td>
</tr>
<tr>
<td>Need to increase opportunities for student internships</td>
<td>0</td>
</tr>
<tr>
<td>Need to improve student full time job placement</td>
<td>2</td>
</tr>
<tr>
<td>Need to increase opportunities for faculty outreach and consulting</td>
<td>0</td>
</tr>
<tr>
<td>Need to support new directions for faculty research</td>
<td>1</td>
</tr>
<tr>
<td>Industry partner company provided significant funding to support program creation</td>
<td>0</td>
</tr>
</tbody>
</table>

Other research points to industry demand playing a role in driving university investment in analytics. The Associate Director of Marketing at the Kelley School of Business (Indiana University) was quoted as saying “Our graduate program in Business Analytics grew out of a direct request from a company who approached us with this need” (Tableau, 2016). According to the Business Analytics Program Director at the Red McCombs School of Business in the University of Texas – Austin, business demand has created a “crushing need for talent” (Tableau, 2016).

These initial findings point to a variety of reasons why colleges and universities may create a data analytics program. Further research is required for conclusive evidence regarding the primary motivating factors in the adoption of data analytics programs.
6. Conclusions and Implications

Data analytics is a very new field of study that incorporates a number of long standing academic areas including math, statistics, computer science, information systems, and application areas in business and the sciences. The programs fill a need in industry and government for people that can analyze vast amounts of data to support organizational decision-making and improve their competitive position. The findings from this study provide some insights into the current practices for data analytics programs in the US regional master’s universities.

Data analytics undergraduate majors and graduate degree programs have almost all been introduced in the past five years. In the sample of universities evaluated in this study, about 25% have a program (98 out of 391). Growth in new programs might slow over the next five years although it can be expected that it is more likely that new programs would come from schools with high overall academic quality, larger enrollments, larger cities, and existing majors in statistics and related quantitative areas. New programs can also be expected to be initiated by university and college administrators and faculty who see a need to attract new students and keep their curriculum up-to-date.

Several directions for future research arise from this study. First, this study’s hypotheses could be addressed using a sample of larger research focused universities or a sample of universities outside the US. Additionally, the motivations for data analytics program development could be addressed in much more detail across a wider range of universities as well as with a larger sample size. This could provide additional insights for educational administrators to identify the best processes to use when developing new technology-related interdisciplinary undergraduate and graduate programs.

The amount of data collected by organizations will continue to grow in the future so the need for data analytics experts can also be expected to continue into the foreseeable future. Meeting this need will require partnerships between universities, businesses, and governmental agencies.
References


Author Biographies

Dr. Troy J. Strader is the Aliber Distinguished Professor of Information Systems at Drake University. He received his Ph.D. in Business Administration (Information Systems) from the University of Illinois at Urbana-Champaign in 1997. His research interests include information technology ethics, digital product management, online consumer behavior, information technology adoption, and the impact of the Internet and e-business on initial public offerings. Dr. Strader has co-edited three books and his work has been published in the *Journal of the Association for Information Systems, European Journal of Information Systems, International Journal of E-Commerce, Communications of the ACM, Decision Support Systems, Electronic Markets* and other academic and practitioner journals and books.

Dr. Andrew Bryant is an Assistant Professor of Marketing at Drake University where he has served since 2014. He received his PhD from the George Washington University School of Business. Andrew’s research interests include Consumer Behavior, Corporate Social Responsibility, and the intersection of Marketing and Public Policy. His work has been published in various forms including journals such as *Business & Society, the Journal of Consumer Affairs, the Journal of Product & and Management, Group & Organizational Management, and Technovation.*