

# Applying Business Analytic Methods To Improve Organizational Performance In The Public School System

Completed Research

**Thomas G. Cech**

Center for Analytics Research and Education  
Appalachian State University  
Boone, North Carolina 28608  
[cechtg@appstate.edu](mailto:cechtg@appstate.edu)

**Trent J. Spaulding, PhD**

Department of Health Care Management  
Appalachian State University  
Boone, North Carolina 28608  
[spauldintj@appstate.edu](mailto:spauldintj@appstate.edu)

**Joseph A. Cazier, PhD**

Center for Analytics Research and Education  
Appalachian State University  
Boone, North Carolina 28608  
[cazierja@appstate.edu](mailto:cazierja@appstate.edu)

## Abstract

This work takes common and innovative business analytics techniques and seeks to apply them in the setting of the school system to improve educational attainment in both primary and secondary schools in the United States. In particular, this paper reviews common factors identified in the literature as influencing a student's success in secondary school, discusses how those factors could be digitized and collected through information systems, and theorizes how big data and analytics could be further applied to these organization to manage their performance. We then look at the uses of analytics in schools to see how well they match and identify areas for improvement. This work shows that there has been an effort to digitize some of the prediction factors; however, a number of the more readily influenced factors have yet to be digitized and used to make evidence based decisions to improve student outcomes in the school system.

**Key Words:** Big Data, Analytics, Educational Attainment, Organizational Performance

## Introduction

Business analytics has produced considerable benefits for organizational performance improvement over the last few years. However, the business sector is no longer the only field that is realizing the benefits of implementing information systems to help improve organizational performance through data-driven decision making. In healthcare, the ability to capture clinical measures relies heavily on being able to capture data through standardized documentation. Successfully implemented electronic health record (EHR) systems allow for healthcare stakeholders to not only improve their practices, but also drastically improve patient outcomes (Weston and Roberts 2013). Unfortunately, few secondary educational institutions have taken advantage of these new technological advances.

Rather than provide all youth with the opportunity to capitalize on their academic abilities, the U.S. education system now tends to unintentionally end up separating students based on their demographic characteristics in a way greatly impacts them to later in life. A study published by the National Center for Education Statistics has found that the national high school graduation rates was at 80%; however, graduation rates of economically disadvantaged students were 70% on average but some states reported rates as low as 53% (Marie and Robert 2014). The gap in outcomes is now larger than it has been in the past. Bachelor's degree attainment rates in 2013 were 66 percentage points lower for students from low-income families than for students from high-income families (Perna 2015). While the socioeconomic factors that bar many students from continuing education cannot be fully addressed with the application of business analytic techniques, other factors such as attendance and grades can. Breiter and Light (2006) argues that the transformation of the public education system into a data-driven enterprise is necessary. This transformation will require new tools, skills, and a culture change to embrace analytics in education.

This paper will argue that the benefits of analytics in education have not been fully realized and that there is a lack of rigorous application of the principles learned from business analytics. The implementation of analytics in education has been slow due several factors. These include, but are not limited to, the emergence of new technology, the educational culture, government mandates that require compliance and a talent shortage. With the newness of big data, educational culture, higher salaries in other sectors for people with big data skills and a need for a rigorous application of the principles for good analytics the education system still has a high degree of growth potential from the application of big data and sophisticated business analytics to improve their organizational performance.

Current attempts to use data are in many cases compliance-focused. They focus on accountability and meeting state and federal requirements more than on systematically investigating the factors that support and hinder the teaching and learning process (Murray 2014). This paper will show that the knowledge from cutting edge businesses that deal with big data can be applied to the public school system and used to solve some of the challenges that these schools are having with managing the educational attainment outcomes for US students. While the data collected and analyzed here may not be as large in scope as is common in business, what this paper proposed is the digitization, collection, and application of analytic methods on data that has traditionally not been used to assist in making evidence based decisions within the context of the public education system. The amount of data available is growing steadily and it is valuable to stay ahead of the data and begin implementing proper IT infrastructure.

In order to examine the current practices in the secondary education system we use Thomas Davenport's definition of analytics. Analytics is "the extensive use of data, statistical or quantitative analysis, explanatory and predictive models, and fact-based management to drive decisions and actions" (Davenport, 2007). In the context of this paper "extensive" suggests using multiple sources of data. "Statistical and quantitative" analyses suggest that, where possible, large amounts of data are used to provide decision support. Finally, interventions are evaluated for the use of data in decision making as opposed to simple compliance and reporting.

This paper starts by identifying key factors that influence a student in reaching their optimal education level. This paper will also examine some current practices with analytics in the school systems and how they can be improved through using the more comprehensive framework outlined above. Next, this paper will take these common factors and explore how they could be digitized and analyzed using big data techniques to both predict when there may be a problem as well as identifying a remedy or possible intervention. Lastly, this paper will conclude with the results of our analysis and application proposals will be discussed.

## **Literature Review**

### ***Factors Identified as Important in the Education Literature***

Similar to Shen and Cooley (2008), we break down the data in to several different perspectives in order to better apply principles of analytics. Based on a review of the literature of the most important factors in keeping students in school, this paper will focus on two distinct types of data: background characteristics

and academic characteristics. We break these two types of data down further into subcategories. These subcategories consist of the individual items commonly identified in the literature as important and are then grouped into larger constructs or categories that can be worked with.

Background characteristics are separated into socioeconomic status, family size and structure, parent characteristics, social engagement, and educational attitudes. The background characteristics data is static, meaning that it does not change on a daily basis. Some of this data is rudimentary and already being gathered while other portions of this data are much more difficult to obtain and will require new and innovative methods of acquisition. These methods will be discussed later in this paper.

Academic characteristics are split into academic performance and school performance. Much of this data is already being gathered through test scores, attendance records, discipline records, and existing school characteristics.

We also separate all factors into two groups. The first group consists of factors that are difficult to influence, but have predictive capabilities. The second consists of those which are predictive and have the potential to be influenced. This separation can be seen below in Tables 1 and 2 respectively.

| <b>Retention Factors</b>  | <b>Sub factors</b>                     | <b>Overall Effects</b>   | <b>Supporting Literature</b>   |
|---------------------------|--|--|--|
| Socioeconomic Status      | Free and reduced lunches               | Students with free or reduced lunches were less likely to succeed in secondary education.  | WWC 2014   |
|                           | Family income                          | Students whose families had low income tended to succeed less often than those with high incomes.  | WWC 2014; Ekstrom et al. 1986; Finn and Rock 1997  |
|                           | Home resources                         | Students with access to many home resources such as books, computers, internet, etc.... tended to be more successful in secondary education. | Finn and Rock 1997; Roderick and Camburn 1999  |
|                           | Region                                 | Those from a more affluent and urban region tended to do better in secondary education as opposed to those from more rural areas.            | Ekstrom et al. 1986  |
| Family Size and Structure | Family Structure                       | Students from two parent households showed better results throughout their secondary educations.   | WWC 2014; Alexander et al. 1997; Ekstrom et al. 1986; Finn and Rock 1997; Rumberger 1995; Rumberger and Larson 1998                                    |
|                           | Teen parent status                     | If the student was a parent then their secondary education suffered.   | WWC 2014; Ekstrom et al. 1986  |
|                           | Number of siblings                     | The more siblings a student has the higher the chance of the student struggling in their secondary education.                                | Alexander et al. 1997; Ekstrom et al. 1986   |
|                           | Special education or disability status | Those students with special education needs or disability status struggled far more than those without.                                      | WWC 2014   |
| Parent Characteristics    | Parent's education                     | The level of education that a student's parent achieved was heavily related to the educational attainment of the student.                    | WWC 2014; Ekstrom et al. 1986; Finn and Rock 1997  |
|                           | Immigrant or English learner status    | If a parent was an immigrant or an English learner the student tended to struggle much more throughout their secondary education.            | WWC 2014; Rumberger 1995   |
|                           | Parents employment                     | Students whose mothers worked tended to struggle in their secondary education.   | Ekstrom et al. 1986; Finn and Rock 1997  |
|                           | Age of mother at birth                 | Students whose mothers were young when they gave birth tended to do poorly during their secondary education.                                 | Alexander et al. 1997  |
| Educational Attitudes     | Parent attitudes                       | Students whose families provided little to no parental support, supervision, or expectations did poorly in their secondary education.        | Alexander et al. 1997; Barrington and Hendricks 1989; Battin-Pearson et al. 1989; Roderick and Camburn 1999; Rumberger 1995; Rumberger and Larson 1998 |
|                           | Student attitudes                      | Students with low self-esteem and self-efficacy tended to do poorly in secondary education.  | Alexander et al. 1997; Battin-Pearson et al. 1989; Finn and Rock 1997; Roderick and Camburn 1999; Rumberger and Larson 1998; Vallerand et al.          |

|                      |                          |  |   |
|----------------------|--------------------------|--|---|
|                      |                          |  | 1997  |
| Academic performance | Standardized Test Scores | Low scores are correlated with poor success in secondary school.                       | Linderman and Baron-Donovan 2006; WWC 2014; Alexander et al. 1997; Finn and Rock 1997; Rumberger and Larson 1998; Jimerson 1999   |
|                      | Grade Point Average      | Low grade point averages were correlated with poor performance in secondary education. | WWC 2014; Alexander et al. 1997; Allensworth and Easton 2005; Allensworth and Easton 2007; Barrington and Hendricks 1989; Roderick and Camburn 1999; Rumberger and Larson 1998; Suh et al. 2007 |
|                      | School Mobility          | The more a student moved the less likely they were to succeed in secondary education   | Finn and Rock 1997; Rumberger 1995; Rumberger and Larson 1998   |

**Table 1. Unsusceptible to Influence Retention Factors**

**Group One – Predictive Factors with Limited Ability to Influence**

The first group of factors (seen in Table 1) consists of those which have been found to predict how a student will progress in their secondary education environment. This group presents factors that have predictive capabilities, but that lack the ability to be directly changed through intervention or policy. The literature has suggested that these sub factors have an impact as to whether students will graduate from their secondary institution; however, the potential for these factors to be influenced through prescriptive data analytics is minimal. Further information about the overall effects of the factors on student retention and academic success are listed in the table.

**Group Two – Predictive Factors with the Potential to be Influenced**

The second group of factors (seen in Table 2) consists of those that have been identified in the literature as being highly correlated with retention factors of students in secondary education. In addition to having predictive qualities, this group of factors has the potential to be influenced to change the overall behavior and outcomes of students. Further information about the overall effect of the factor on student retention and academic success is found in the table.

| Retention Factors    | Sub factors   | Overall Effects  | Supporting Literature  |
|----------------------|---|--|--|
| Social Engagement    | Burn out  | Students with high educational expectations but few resources will tend to do poorly throughout their secondary education.       | Salmela-Aro and Upadyaya 2014  |
|                      | Working   | Students who work 10+ hours a week tend to do better in secondary education while those who work 10+ per week tend to do poorly. | Marsh and Kleitman 2005  |
|                      | Extracurricular participation                           | Students who participate in extracurricular activities tend to do better in their secondary education.                           | Finn and Rock 1997; Rumberger and Larson 1998  |
| Academic Performance | Rate of school attendance                               | Students who attend classes more regularly tend to do better.  | WWC 2014; Allensworth and Easton 2007; Barrington and Hendricks 1989; Battin-Pearson et al. 1989; Rumberger 1995; Rumberger and Larson 1998; Jimerson 1999 |
|                      | Frequency of behavior or discipline incidents in school | Those with fewer disciplinary or behavioral incidents succeed more often in secondary education.                                 | WWC 2014; Battin-Pearson et al. 1989; Finn and Rock 1997; Rumberger and Larson 1998; Suh et al. 2007; Jimerson 1999  |
|                      | Performance at  | Students who are behind in their grade level have a  | WWC 2014; Alexander et al.   |

|                                |   |   |   |
|--------------------------------|---|---|---|
|                                | Grade level   | difficult time catching up and as a result tend to do poorly in secondary education.                                | 1997; Allensworth and Easton 2005; Allensworth and Easton 2007; Barrington and Hendricks 1989                     |
| Academic Performance Continued | Classes schedule  | Students enrolled in more difficult classes tend to do better in their secondary education.                         | Norton 2011   |
|                                | Classes for college credit  | Students who are enrolled in concurrent classes or AP classes succeed more in their secondary education.            | National Center for Educational Statistics 2014   |
| School Performance             | Guidance feedback   | Students that attend schools with access to guidance and feedback from counselors tend to do better.                | Solberg et al. 2007; Hovdhaugen et al. 2013; Norton 2011  |
|                                | Facility resources  | Students that attend schools with current technology as well as access to counselors and mentors tend to do better. | Hovdhaugen et al. 2013; Solberg et al. 2007; Kalsbeek and Zucker 2013; Norton 2011; Salmela-Aro and Upadyaya 2014 |
|                                | School characteristics (Staff ability levels and school demographics) | Students who attend schools with low teacher/student ratios, crime, and numbers of students tend to do better.      | Finn and Rock 1997  |

**Table 2. Predictive and Directly Influenceable Retention Factors**

## Uses of Analytics in the Public Education System

As discussed above, the factors identified through the literature review have the potential to predict if a student is having a suboptimal education experience. Having separated these predictive factors above, this paper will now look at how these factors are being used in the literature today. According to Thomas Davenport’s definition of analytics, the literature suggests that there is some extensive use of analytics implemented throughout the education system, but there is still much more that can be done to improve the educational outcomes of students.

The last few years have seen dramatic increases in the ability to predict which students are at a high risk of dropping out. Through the use of the above-mentioned factors, the predictions are at the point where we can intervene at both the district level as well as the individual level; however, the data type used is generally the unsusceptible data that is discussed above. The use of these factors provides schools with relatively accurate predictions as to which students will require interventions. The majority of interventions that have been suggested through the literature rely on the results of standardized tests and grade point averages.

These factors, though important, are difficult to influence through policy or intervention. Others include more easily influenced academic performance metrics such as attendance and frequency of behavior incidents, but most tracked factors are related to socioeconomic status, grade point average, and standardized testing. As a result, these tracked retention factors are rarely used to drive decision making throughout the educational system. Instead, these factors are used for accountability and local and federal reporting. Lastly, most prediction is based on single retention factors rather than several. Very few interventions are focused on influencing the factors that are found in the second group of directly influenced factors.

Table 3 is a more comprehensive overview of the various factors that are being used throughout the literature to prescribe various interventions. Efforts need to focus more on addressing the factors that can be directly influenced rather than those that are unchangeable in the short term. The explosion of interest in this area has allowed for many innovations in predicting which students will require help; however, little has been done to prescribe and prioritize which interventions will be most applicable and helpful to an individual student throughout their academic career.

There are large gaps in the data that the literature has identified as valuable for predicting student outcome, but has been rarely digitized or used. An analysis of a combination of factors from group one

and group two would allow for the proper allocation of limited resources based on prescriptive analytic methods. How the more easily influenced data could be used is further explored below in the discussion section.

| Retention Factors          | Intervention Treatment                                       | Extent of Data Use | Statistical Analysis | Model Type  | Data Driven Decisions | Supporting Literature  |
|----------------------------|--|--------------------|----------------------|-------------|-----------------------|--|
| Academic Performance Data  | Tutoring and academic enrichment                             | Extensive          | Yes                  | Explanatory | Yes                   | Linderman and Baron-Donovan 2006; Yampolskaya et al. 2006; Ward et al. 2013; Forum for Youth Investment 2014           |
|                            | Summer enrichment programs                                   | Extensive          | Yes                  | Predictive  | Yes                   | Linderman and Baron-Donovan 2006   |
|                            | Computer skills training                                     | None               | No                   | None        | No                    | Linderman and Baron-Donovan 2006   |
|                            | Mentoring  | Extensive          | Yes                  | Explanatory | No                    | Mthethwa-Sommers 2013; Monk et al. 2014  |
|                            | Counseling, advising, career planning, and academic planning | Extensive          | Yes                  | Predictive  | Yes                   | Linderman and Baron-Donovan 2006; Yampolskaya et al. 2006; Camizzi et al. 2009; Ward et al. 2013; Tillery 2013         |
|                            | College visits   | Extensive          | Yes                  | Explanatory | No                    | Linderman and Baron-Donovan 2006; Tillery 2013; Mthethwa-Sommers 2013; Ward et al. 2013                                |
|                            | Materials development and other                              | Minimal            | No                   | None        | No                    | Linderman and Baron-Donovan 2006   |
|                            | Educational field trips                                      | Minimal            | No                   | None        | No                    | Linderman and Baron-Donovan 2006   |
|                            | Workshops  | Minimal            | Yes                  | Explanatory | No                    | Linderman and Baron-Donovan 2006; Yampolskaya et al. 2006; Weiher et al. 2006; Mthethwa-Sommers 2013; Ward et al. 2013 |
|                            | Job site visits and shadowing                                | Minimal            | Yes                  | Explanatory | No                    | Linderman and Baron-Donovan 2006   |
| Educational Attitudes Data | Student Family events  | Minimal            | No                   | None        | No                    | Linderman and Baron-Donovan 2006; Castillo et al. 2010; Tillery 2013   |
|                            | Student Cultural Events                                      | Minimal            | No                   | None        | No                    | Linderman and Baron-Donovan 2006; Tillery 2013   |
|                            | Parent Family Events   | Minimal            | Yes                  | Explanatory | No                    | Tillery 2013   |
|                            | Parent Cultural Events                                       | Minimal            | Yes                  | Explanatory | No                    | Gibson 2006; Tillery 2013  |
|                            | Parent Counseling/ advising                                  | None               | No                   | None        | No                    | Tillery 2013   |
|                            | Parent Workshops   | Extensive          | Yes                  | Explanatory | No                    | Davis 2000; Tillery 2013   |
| School Performance Data    | Enhancing the teachers qualification                         | Extensive          | Yes                  | Explanatory | No                    | Linderman and Baron-Donovan 2006; Castillo et al. 2010; Kim 2010; Murray 2014  |

**Table 3. Use of Data for Analytics in the Public Education System**

In discussing data analytics and the potential uses of the above identified factors, it is important to be able to distinguish between the various types of analytics and how they can be applied to the data. Table 4 presents the factors that have been identified in the literature above and highlights the current levels of

use in analytics and what future applications might look like. In doing so, this work discusses four types (descriptive, diagnostic, predictive, and prescriptive) of analytics commonly employed.

| Retention Factors         | Sub factors   | Current Level of Use in Public Schools | Future Analytic Applications <sup>1</sup> |
|---------------------------|---|--|---|
| Socioeconomic Status      | Free and reduced lunches                                | Medium - High                          | Predictive                                |
|                           | Family income   | Medium                                 | Predictive                                |
|                           | Home resources  | Low                                    | Predictive                                |
|                           | Region  | Low                                    | Descriptive/Diagnostic                    |
| Family Size and Structure | Family Structure  | Low                                    | Descriptive/Diagnostic                    |
|                           | Teen parent status                                      | Low                                    | Predictive                                |
|                           | Number of siblings                                      | Low                                    | Descriptive/Diagnostic                    |
|                           | Special education or disability status                  | Medium                                 | Predictive                                |
| Parent Characteristics    | Parent's education                                      | Low - Medium                           | Predictive                                |
|                           | Immigrant or English learner status                     | Low                                    | Descriptive/Diagnostic                    |
|                           | Parents employment                                      | None                                   | Descriptive/Diagnostic                    |
|                           | Age of mother at birth                                  | None                                   | Descriptive/Diagnostic                    |
| Educational Attitudes     | Parent attitudes  | Medium                                 | Predictive/Prescriptive                   |
|                           | Student attitudes                                       | High                                   | Predictive/Prescriptive                   |
| Academic performance      | Standardized test scores                                | High                                   | Predictive                                |
|                           | Grade point average                                     | High                                   | Predictive                                |
|                           | School mobility   | Low                                    | Descriptive/Diagnostic                    |
| Social Engagement         | Burn out  | None                                   | Predictive/Prescriptive                   |
|                           | Working   | None                                   | Predictive                                |
|                           | Extracurricular participation                           | Low - Medium                           | Predictive/Prescriptive                   |
| Academic Performance      | Rate of school attendance                               | High                                   | Predictive/Prescriptive                   |
|                           | Frequency of behavior or discipline incidents in school | High                                   | Predictive                                |
|                           | Performance at Grade level                              | Low - Medium                           | Predictive/Prescriptive                   |
|                           | Classes schedule  | None                                   | Predictive/Prescriptive                   |
|                           | Classes for college credit                              | None                                   | Predictive/Prescriptive                   |
| School Performance        | Guidance feedback                                       | Low - Medium                           | Predictive/Prescriptive                   |
|                           | Facility resources                                      | Low                                    | Prescriptive                              |
|                           | School characteristics                                  | Low                                    | Predictive/Prescriptive                   |
|                           | School demographics                                     | Medium- High                           | Predictive                                |

**Table 4. Current and Future Uses of Retention Factors for Analytics**

Descriptive analytics focuses on what has happened in the past and what is happening now. Diagnostic analytics are concerned with why a particular event occurred. Descriptive analytics and diagnostic analytics are often used in tandem to examine key performance indicators and key metrics in order to understand an organizations performance and to evaluate an organizations process. Predictive analytics are concerned with predicting what is likely to happen in the future. This analytic process often employs predictive models to identify patterns in the available data. The prescriptive analytic process also employs predictive models as well as optimization techniques to recommend one or more courses of action.

<sup>1</sup> The potential value of the factors presented depends on the perspectives of the stakeholders and policymakers.

Though all analytic processes are valuable in their own way, this last analytic process holds the most potential for improving student outcomes in the public education system.

Table 4 provides an overview of the current state of the use of analytics in the public school system on each factor identified as important as well as the potential use in the future on the realm of descriptive, diagnostic, predictive and prescriptive.

Results show that while some level of analytics is being used in the public school system, many of the previously identified factors deemed important are currently not being analyzed or are being analyzed at a very low level. Additionally there appears to be much potential for future usage to be explored in future research that could add value to the overall organizational performance through the strategic use of business analytics. The factors discussed in table 4 reflect those that have been identified in tables 1 and 2 as either more difficult or easier to influence by policy makers. The factors ability to be influenced affects whether or not analytic methods could be applied for prescriptive interventions. The factors identifies are those which are currently used to predict outcomes in the secondary public school education system.

## **Discussion**

In order to correctly apply business analytic methods to enhance the organizational performance within the domain of the public education system, several steps need to take place. First, a move for further digitization and better collection of the current data that has traditionally been used within the public education system as well as novel ways of collecting data which are atypical within the context of education. Second, the application of analytic methods to improve organizational performance, which have traditionally not been used within the public education system in order to help policy makers and primary stakeholders to make evidence based decisions. Lastly, to provide the knowledge gained through the application of these methods in the form of key performance indicators back to the primary stakeholders (the states, districts, schools, teachers, parents, and students).

In addition to the use of academic performance data, it would be valuable to begin digitizing and tracking several of the factors that are found in the second group this paper has identified. Specifically, these would include the social engagement factors and the school performance factors. These two factors tend to be overlooked in the literature, but could provide valuable information to help drive future evidence based decisions in the public education system. The school performance factors are already being captured and digitized but, again, are only used for recording and accountability.

Social engagement data is one of the least tracked retention factors, but shows the greatest opportunity in learning about how engaged a student is within the community. A student's engagement within the community often provides information about their susceptibility to join gangs or to participate in drug use. Those students who are engaged more within the community are less likely to associate themselves with these anti-social behaviors. Having access to this information would allow the public education system to implement interventions and policies that can directly impact these students.

Academic performance metrics measuring the difficulty of classes a student is enrolled. Often students find themselves enrolled in classes that are either too easy or difficult for their academic abilities. The ability to tailor a student's course load to match their academic ability would allow students to be more confident in the classes that they are taking and give them the opportunity to focus on the areas that they need help in.

School performance data has the potential to inform public school policy makers about the various environmental factors that play into a student's overall educational attainment. Policies and interventions can be designed to improve the overall educational environment in which students learn in.

Much more needs to be done to capture this valuable predictive data, and use it for prescriptive purposes. In order for these prescriptive analytics to start taking place in the educational system, more digitization and new analytic techniques need to take hold. Traditionally, there have been many issues with gathering information that captures a student's social engagement and self-esteem. Documents such as diaries



autobiographies, and/or letters inform researchers about the students' self-esteem and social engagement. Diaries seem to constitute a material in which people write regularly about their thoughts and actions, without the awareness of an audience. Zittoun and Gillespie suggest that diaries – and to some extent, other forms of self-writing – might offer an important source of data for the scholar interested in human development.

Analytical and theoretical tools developed over the past few years to analyze thinking, can be used to analyze diaries. As a result, diaries and self-writing can be said to offer a unique entry into the analysis of sociocultural and developmental processes (Zittoun and Gillespie 2012). Digitizing a student's portfolio of self-writing would allow analysts and researchers to use new methods of analysis – such as text and sentiment analysis – to identify the above mentioned psychological, social, and opinion related factors. The implementation of text and sentiment analysis will allow teachers and other stakeholders to identify the various interventions that may benefit an individual student to stay in school and continuing their education.

One of the more difficult factors to identify for many schools is a student's socioeconomic status. While many measurements exist and are tracked, access to these factors is difficult to obtain. However, schools do have access to a student's address and this could be used to tie the student to a particular socioeconomic status. This is just one use of a student's home address. Identifying where a student lives in relation to a school has the potential for schools to identify why students are absent, tardy, or overly exhausted. Many students (especially in rural areas) can spend extended period of times en route to and from school. Through the application of analytic techniques, evidence based interventions could be provided to optimize transportation systems to and from school.

Other factors such as parent attitudes towards education could be gathered through proxies. These proxies could include the parental involvement within school, parent education, and parent surveys that are distributed online. Student attitudes and perseverance could also be collected in surveys that are designed to measure these specific factors. Surveys (such as the GRIT scale) that help identify these factors already exist and are distributed among students but are not digitized in their current state.

All of the important stakeholders (states, districts, schools, teachers, and families) have had access to traditional sources of data but it has been a challenge for them to make sense of volume of data presented to them. The application of business analytic methods would help provide a clear picture of what needs to be addressed in each state, district, school, and classroom. In the same way that Blue Cross Blue Shield uses data to not only make their organization more efficient but also to provide better service to their customers, the public education system has the opportunity to create a more efficient and effective organization.

Schools currently use student portals to inform students of their grades and attendance but only at the class level. Students can already upload written assignment as well as complete online quizzes through these portals, but the digitized versions of these assignments contain data that is not being utilized beyond the teacher grading them. These online submissions are the beginnings of the proposed student portfolios; however, schools, counselors, and teachers have yet to apply current analytic methods to glean further helpful decision making information from them.

In the same way that businesses use many sources of data to indicate the overall performance of the organization, public school systems have access to several sources of data (the factors mentioned above, online student portfolios, and student addresses – just to mention a few of the traditional sources) that can help identify how the organization is operating at different levels. Once all of the traditional data as well as some of the above proposed data sources have been digitized then states, districts, schools, and teachers would have a deeper understanding of what the strengths and weaknesses of each district, school, teacher, and student that they are responsible for. Having the ability to identify the strengths and weaknesses through key performance indicators would better enable the stakeholders to make informed decisions based on evidence provided through the analysis of all available digitized sources.

## Conclusion

This paper has identified specific factors that could predict and improve educational attainment in the public education system. By digitizing and analyzing social engagement and school performance factors, student outcome could be greatly improved. The literature review reveals that many factors used today to predict student outcomes are difficult to directly influence. Analytic methods in the education system have made good use of the academic performance factors of attendance and frequency of behavioral problems, but have not taken into account some of the others that have potential. The factors discussed here are those that have the most potential to make evidence based decisions; however, with further advances in technology and analytic methods there are bound to be others that arise and have not been addressed in this paper.

Public education interventions have come a long way, but until now have only focused on explaining and predicting which students may be at risk of failing or not continuing their education; however, these interventions fail to make use of the data that is available to them by not providing prescriptive solutions to the issues that at-risk students are faced with. This paper suggests that data analytics has the potential to not only help those that are at risk of failing or dropping out of school, but also those that are at the top of their class. Customizing education will enable public education organizations to meet the needs of every individual student, rather than simply focusing on saving those students that are the most challenged. This paper does not suggest that the implementation of analytics will be perfect; it argues that evidence based decisions are an improvement upon the system that is currently in place.

The future application of business analytic techniques has the potential to transform the public education system to a more optimized organization that is able to identify strengths and weaknesses. In its current state the data gathered and analyzed may not be as large, varied, or continuous as it is in businesses, but with future exploration and application these data sources would be as important in making the organization as effective and efficient as it is within the business domain.

## References

- Alexander, K., Entwisle, D., & Horsey, C. (1997). From first grade forward: early foundations of high school dropout. *Sociology of Education*, 70, 87–107.
- Allensworth, E. M., & Easton, J. Q. (2005). The on track indicator as a predictor of high school graduation. Chicago, IL: Consortium on Chicago school research. Retrieved 3 February 2015, from <http://ccsr.uchicago.edu/publications/p78.pdf>
- Allensworth, E. M., & Easton, J. Q. (2007). What matters for staying on track and graduating in Chicago Public High Schools. Chicago, IL: Consortium on Chicago school research. Retrieved 3 February 2015, from [http://ccsr.uchicago.edu/content/publications.php?pub\\_id=116](http://ccsr.uchicago.edu/content/publications.php?pub_id=116)
- Barrington, B. L., & Hendricks, B. (1989). Differentiating characteristics of high school graduates, dropouts, and nongraduates. *The Journal of Educational Research*, 82(6), 309–19.
- Battin-Pearson, S., Newcomb, M. D., Abbott, R. D., Hill, K. G., Catalano, R. F., & Hawkins, J. D. (2000). Predictors of early high school dropout: A test of five theories. *Journal of Educational Psychology*, 92(3), 568–82.
- Breiter, A., & Light, D. (2006). Data for School Improvement: Factors for designing effective information systems to support decision-making in schools. *Journal Of Educational Technology & Society*, 9(3), 206-217.
- Camizzi, E., Clark, M., Yacco, S., & Goodman, W. (2009). Becoming "Difference Makers": School-University Collaboration to Create, Implement, and Evaluate Data-Driven Counseling Interventions. *Professional School Counseling*, 12(6), 471-479.
- Castillo, L. I., Conoley, C. W., Cepeda, L. M., Ivy, K. K., & Archuleta, D. J. (2010). Mexican American Adolescents' Perceptions of a Pro-College Culture. *Journal Of Hispanic Higher Education*, 9(1), 61-72. doi:10.1177/1538192709350454
- Davenport, T., & Harris, J. (2007). *Competing on analytics: The new science of winning*. Boston, Mass.: Harvard Business School Press.
- Davis, D. (2000). Supporting Parent, Family, and Community Involvement in Your School. *Northwest Regional Educational Laboratory*. Retrieved February 18, 2015.

- Ekstrom, R. B., Goertz, M. E., Pollack, J. M., & Rock, D. A. (1986). Who drops out of high school and why? Findings from a national study. *Teachers College Record*, 87, 356–73.
- Finn, J. D., & Rock, D. A. (1997). Academic success among students at risk for school failure. *Journal of Applied Psychology* 82(2), 221–34.
- Forum for Youth Investment. (2014, July). Sharing data is a calculated move: Schools partner with afterschool providers to improve youth services. Retrieved January 17, 2015, from [http://www.readyby21.org/sites/default/files/Nashville\\_Sharing\\_Data\\_July2014.pdf](http://www.readyby21.org/sites/default/files/Nashville_Sharing_Data_July2014.pdf).
- Gibson, D., & Jefferson, R. (2006). The effect of perceived parental involvement and the use of growth-fostering relationships on self-concept in adolescents participating in GEAR UP. *Adolescence*, 41(161), 111-125.
- Hovdhaugen, E., Frolich, N., & Aamodt, P. O. (2013). Informing Institutional management: Institutional strategies and student retention. *European Journal of Education*, 84, 165-177. doi: 10.1111/ejed.12002
- Jimerson, S.R. (1999). On the failure of failure: Examining the association between early grade retention and education and employment outcomes during late adolescence. *Journal of School Psychology*, 37, 243–272.
- Kalsbeek, D. H., & Zucker, B. (2013). Reframing retention strategy: A focus on profile. *New Directions for Higher Education*, 161, 15-25. doi: 10.1002/he.20042
- Kim, T. (2010). An Effective Way to Improve Mathematics Achievement in Urban Schools. *Educational Research Quarterly*, 34(2), 60-71.
- Linderman, D., & Baron-Donovan, C. (2006, June). Collaborative Programs Research Report: CUNY GEAR UP. Retrieved February 10, 2015, from [http://www.cuny.edu/academics/evaluation/library/cn\\_gearup\\_donovan.pdf](http://www.cuny.edu/academics/evaluation/library/cn_gearup_donovan.pdf)
- Marie C., S., & Robert, S. (2014). Public High School Four-Year OnTime Graduation Rates and Event Dropout Rates: School Years 2010–11 and 2011–12. *National Center for Education Statistics*. Retrieved April 1, 2015, from <http://nces.ed.gov/pubs2014/2014391.pdf>
- Marsh, H. W, and Kleitman, S. (2005) Consequences of employment during high school : Character building, subversion of academic goals, or a threshold? *American Educational Research Journal* 42: 2. 331-369.
- Monk, M. m., Baustian, M. M., Saari, C. R., Welsh, S., D'Elia, C. F., Powers, J. E., & ... Francis, P. (2014). EnvironMentors: Mentoring At-Risk High School Students through University Partnerships. *International Journal Of Environmental & Science Education*, 9(4), 385-397. doi:10.12973/ijese.2014.223a
- Mthethwa-Sommers, S. s. (2013). Reading Beyond Research Results: The Hidden Curriculum in a College and Urban High Schools Partnership. *Journal Of Community Engagement & Higher Education*, 5(2), 45-55.
- Murray, J. (2014). Critical Issues Facing School Leaders Concerning Data-Informed Decision-Making. *Professional Educator*, 38(1), 14-22.
- National Center for Education Statistics. (2014, April 25). HSLs:09 – 2013 updated instrument. Retrieved from [http://nces.ed.gov/surveys/hsls09/pdf/2013\\_Student\\_Parent\\_Questionnaire.pdf](http://nces.ed.gov/surveys/hsls09/pdf/2013_Student_Parent_Questionnaire.pdf)
- Norton, M. S. (2011). Please, not another push to get touch on student retention. *Planning and Changing*, 42, 209-223.
- Perna, L. (2015). *Indicators of higher education equity in the United States*. [Washington]: Pell Institute for the Study of Opportunity in Higher Education. [http://www.pellinstitute.org/downloads/publications-Indicators\\_of\\_Higher\\_Education\\_Equity\\_in\\_the\\_US\\_45\\_Year\\_Trend\\_Report.pdf](http://www.pellinstitute.org/downloads/publications-Indicators_of_Higher_Education_Equity_in_the_US_45_Year_Trend_Report.pdf).
- Roderick, M., & Camburn, E. (1999). Risk and recovery from course failure in the early years of high school. *American Educational Research Journal*, 36, 303–43.
- Rumberger, R. W. (1995). Dropping out of middle school: a multilevel analysis of students and schools. *American Educational Research Journal*, 32, 583–625.
- Rumberger, R. W., & Larson, K. A. (1998). Student mobility and the increased risk of high school drop out. *American Journal of Education*, 107, 1–35.
- Salmela-Aro, K., Upadyaya, K. (2014). School burnout and engagement in the context of demands-resource model. *British Journal of Educational Psychology*, 84, 137-151. doi: 10.1111/bjep.12018
- Shen, J., & Cooley, V. E. (2008). Critical issues in using data for decision-making. *International Journal of Leadership in Education*, 11(3), 319–329.

- Solberg, V. S. H., Carlstrom, A. H., Howard, K. A. S., & Jones, J. E. (2007). Classifying at-risk high school youth: The influence of exposure to community violence and protective factors on academic and health outcomes. *The Career Development Quarterly*, 55, 313-327. doi: 10.1002/j.2161-0045.2007.tb00086.x
- Suh, S., Suh, J., & Houston, I. (2007). Predictors of categorical at-risk high school dropouts. *Journal of Counseling and Development*.
- Tillery, C. (2013). The Summative Impact of College Access Interventions: A Program Evaluation of GEAR UP North Carolina.
- Vallerand, R. J., Fortier, M. S., & Guay, F. (1997). Self-determination and persistence in a real-life setting: toward a motivational model of high school dropout. *Journal of Personality and Social Psychology*, 72, 1161-76.
- Ward, N. n., Strambler, M. J., & Linke, L. H. (2013). Increasing Educational Attainment among Urban Minority Youth: A Model of University, School, and Community Partnerships. *Journal of Negro Education*, 82(3), 312-325.
- Weihner, G. R., Hughes, C., Kaplan, N., & Howard, J. Y. (2006). Hispanic College Attendance and the State of Texas GEAR UP Program. *Review Of Policy Research*, 23(5), 1035-1051. doi:10.1111/j.1541-1338.2006.00248.x
- Weston, M., & Roberts, D. (2013). The Influence of Quality Improvement Efforts on Patient Outcomes And Nursing Work: A Perspective from Chief Nursing Officers at Three Large Health Systems. *The Online Journal of Issues in Nursing*, 18(3).
- What Works Clearinghouse (March 2014). WWC Evidence Review Protocol For Dropout Prevention Interventions, Version 3.0. Retrieved from [http://ies.ed.gov/ncee/wwc/pdf/reference\\_resources/wwc\\_dp\\_protocol\\_v3.0.pdf](http://ies.ed.gov/ncee/wwc/pdf/reference_resources/wwc_dp_protocol_v3.0.pdf)
- Yampolskaya, S., Massey, O. T., & Greenbaum, P. E. (2006). At-Risk High School Students in the 'Gaining Early Awareness and Readiness Program' (GEAR UP): Academic and Behavioral Outcomes. *The Journal Of Primary Prevention*, 27(5), 457-475. doi:10.1007/s10935-006-0050-z0
- Zittoun, T., & Gillespie, A. (2012). Using Diaries and Self-writings as Data in Psychological Research. *Emerging Methods in Psychology*, 1-26. Retrieved February 22, 2015, from [http://www.academia.edu/1308394/Using\\_diaries\\_and\\_self-writings\\_as\\_data\\_in\\_psychological\\_research](http://www.academia.edu/1308394/Using_diaries_and_self-writings_as_data_in_psychological_research)