STRATEGIC DATA PROJECT

SDP FELLOWSHIP CAPSTONE REPORT

Meta Matters: Leveraging Metadata to Improve Data Use and Effectiveness

Lisa Deffendall, Fayette County Public Schools
Lu Han, Syracuse Public Schools
William R. Buchanan, Mississippi Department of Education

SDP Cohort 5 Fellows

Strategic Data Project (SDP) Fellowship Capstone Reports

SDP Fellows compose capstone reports to reflect the work that they led in their education agencies during the two-year program. The reports demonstrate both the impact fellows make and the role of SDP in supporting their growth as data strategists. Additionally, they provide recommendations to their host agency and will serve as guides to other agencies, future fellows, and researchers seeking to do similar work. *The views or opinions expressed in this report are those of the authors and do not necessarily reflect the views or position of the Center for Education Policy Research at Harvard University.*

Abstract

Local education agencies (LEAs) and state education agencies (SEAs) routinely invest significant funds in tools and technologies to facilitate data (Topol, Olson, Roeber, & Hennon, 2012). However, the field of education lags behind the technology sector with regards to leveraging data mining techniques to gain deeper insight into user experience, usage metrics, how these data are related to outcome metrics and how to apply methods appropriate in the context of education (Baker & Yacef, 2009). We used ~39,000 views of NWEA (Northwest Evaluation Association) reports from Fayette County Public Schools to classify educators into discrete latent classes using multilevel latent class analysis. Our findings suggest there are five distinct user groups and that these groups are relatively invariant to factors such as educational level taught and the total number of days on which the platform was accessed. Additionally, we also show initial evidence of relationships between the school-level aggregated frequency of users of a given latent class and changes in accountability system metrics.

The use of student data systems to improve education and help students succeed is a national priority (Means, Chen, DeBarger, & Padilla, 2011; Means, Padilla, & Gallagher, 2010). Data can inform educators about their decision making at all levels which eventually can improve student achievement (Datnow, Park, & Wohlstetter, 2007; Hamilton et al., 2009; Lachat & Smith, 2005; Wayman & Stringfield, 2006). Thus, schools, districts, and state education agencies and institutions invest significant resources annually on tools intended to help make better decisions, such as data dashboards, early warning systems, formative and/or benchmark assessments. According to a report by IDC Government Insights, IT spending by K–12 in the United States is expected to hit about $4.7 billion for 2015 and the expenditure will grow at a constant, steady pace (Topol, Olson, Roeber, & Hennon, 2012). Despite the increased spending on information technology in K–12 education, the amount of research demonstrating how the technology yield a reasonable return on investment is sparse at best. In this report, we attempt to find answers to the following questions:

- How does one know whether data tools in education are used effectively and efficiently?
- Who uses the tools and how do the users interact with the data tools?
- Does the use of these tools have a direct or indirect positive influence on student outcomes?

We approach these answers to these questions by analyzing metadata. Metadata are the data—or information—about a given datum or collection of data. We focused our efforts on analyzing metadata—specifically the server's log files—from the Northwest Evaluation Association's (NWEA) Measures of Academic Progress (MAP) online reporting tool. Our exploratory data analyses approach of these data relied on a free and open source software stack, the Elasticsearch, Logstash, and Kibana or ELK stack. We then moved on to analytical methods designed to help us simplify our understanding of these complex behaviors by classifying—or categorizing—users into discrete groups based on their use of the data tool. We hope to build upon emerging efforts to use data mining, machine learning, and data science techniques in the K–12 educational context to help educational leaders better understand the *types* of data users and possible implications that arise when the users are *too homogeneous*.

Review of the Literature

Educational Data Mining

Data mining also known as "knowledge discovery in database" (KDD) is a series of data analysis techniques applied to extract hidden knowledge from raw data (Whitten & Frank, 1999 as cited in Baker & Yacef, 2009) using a combination of exploratory data analysis, pattern discovery, and predictive modeling (Panov, Soldatova, & Dzeroski, 2009). Data mining continues a history of adoption and acceptance in industries such as business and commerce, healthcare, and technology, but the adoption of these techniques in the education sector is still in its infancy. However, as Baker and Yacef (2009) point out, data mining in the context of education is different for several important reasons that require analysts to address the lack of independence of observations (e.g., students clustered within classrooms, clustered within schools, clustered within districts, etc…) and the use and incorporation of psychometric models used to estimate relationships among characteristics that are not directly observable (e.g., ability, skill, etc…).

Romero & Ventura (2010) reviewed 306 articles from 1993 to 2009 regarding educational data mining (EDM) and proposed desired EDM objectives based on the roles of users. They summarized eleven objectives of EDM research works:

1. analysis and visualization of data (35 research);
2. providing feedback for supporting instruction (40 research);
3. recommendations for students (37 research);
4. predicting students' performance (76 research);
5. student modeling (28 research);
6. detecting undesirable student behaviors (23 research);
7. grouping students (26 research);
8. social network analysis (15 research);
9. developing concept maps (10 research);
10. constructing courseware (9 work);
11. planning and scheduling (11 research).

Another meta-analysis conducted by Peña-Ayala (2014) reveals an EDM work profile was compiled to describe 222 EDM approaches and 18 tools. By the end of the study, the author concludes: ‘‘EDM is living its spring time and preparing for a hot summer season.’’

Data Mining Applications

To date, many applications of data mining techniques in education are targeted at student learning applications (Baker, Corbett, & Gowda, 2013; Baker & Corbett, 2014; Ocumpaugh, Baker, & Rodrigo, 2015). One notable exception to this is an evaluation of the Achievement Reporting and Innovation System of New York City Schools (Gold et al., 2012).

Gold et al. (2012) evaluated usage metrics of a data reporting system in an attempt to answer broader surface questions related to whether/if the tool was used, by whom it was used, and how users interacted with the system. However, by aggregating usage data by user and eliminating sessions lasting more than an hour, the authors' analyses fail to account for the time dependence between the usage/activities within the system (e.g., amount of time elapsed between viewing different reports, or sequential effects). It does, however, represent a major step forward from Baker and Yacef's (2009) description of the infancy of the field only three years prior.

Data Use and Student Outcomes. Tyler and McNamara's (2011) work provides a framework upon which investigations into the effect of educators' data use on student learning outcomes could be estimated. After manually cleaning and parsing server log files, the authors attempted to estimate the effect of the usage of a district-wide implementation of a data analysis/reporting tool. Although use of the tool was encouraged, the degree of uptake was likely insufficient to determine any conclusive dosage effect. In other words, the tool was not used frequently and consistently enough by the educators for the purpose of estimating the returns on a unit increase of data use. This, however, still does not address issues potentially related to the timing of the dosage and the assumptions imposed on the functional form of the model under different designs for time-dependent measurement(s) (Little, 2013). For example, is it reasonable to assume that viewing the data of a student the day prior to school beginning has the same effect on student outcomes as viewing the data for the same amount of time after the mid-term? Would viewing data for a given student once have the same effect as viewing data on a given student multiple times? Are there any pathway effects related to the order/sequence of the reports viewed by an educator? And should the data be treated as purely independent measures at multiple points in time?

Another critical study investigated an innovative approach of program evaluation through analyses of student learning logs, demographic data, and end-of-course evaluation surveys in an online K–12 supplemental program and proposed a program evaluation decision making model based on educational data mining (Hung, Hsu, & Rice, 2012).

Analysis

Study Sample

Observations of 39,925 NWEA Map report views from 3,865 of the 5,182 staff members in the Fayette County Public Schools (FCPS) were collected between November 30th, 2014, 16:36:19 until May 20th, 2015, 13:08:47 and used as the foundation for our research; users had between 0 and 416 report views during the period (mean = 27.43, standard deviation = 36.86). We also implemented some additional post-hoc exploratory research focused on the relationship between concentration of specific user types in a school and educational outcomes. We used school level data from the educational accountability system implemented in Kentucky to provide aggregate measures of student learning at the school level. In other words, we begin by classifying who and how users leverage these data and then move on to exploring how the combination of educators from these different groups is related to student learning measures.

Demographics. FCPS serves approximately 40,000 students across 60+ schools and special programs; the LEA serves a diverse community in which nearly 54% of students qualify for free or reduced-price meals, 54.3% of students are White, 22.6% of students are African

American, 14.3% of students are Hispanic, 4.2% of students are Asian, and 3,789 students are classified as having limited English proficiency (LEP). The sample included 317 unique job titles for the individuals in the data set, which were grouped into the seven categories listed in Table 1 below. The classroom educator category was further disaggregated in an attempt to separate and identify grade spans associated with the educational services provided (see Table 2 for additional information).

What are we classifying? We analyzed data dashboard log files that included unique educator identifiers, timestamps, and an indicator of what type of report the educator was viewing. In addition, we joined official job titles to these data to use for subsequent analysis to determine whether or not job classification, average number of reports viewed each day the user had activity, and the total number of days that users accessed the platform predicted the classification of the users; for example, would being a school administrator increase the likelihood that they would be classified into one of the user classes? In the end, the goal is to identify distinct groups, or classes, of educators based on their interaction with the data tool and use these insights to support decisions regarding reinvestment in analytical platforms, crafting professional development, and/or informing decisions about professional learning community composition (e.g., to ensure that each PLC has at least one strong data user). We show how report views vary by job type/function for all of the records that were analyzed in table 3.

Methods

Given our primary goal of understanding the types of educational data system users, we wanted to organize these discussions around discrete groups of users. Although several models exist for classification problems in the context of supervised and/or semi-supervised (Information Resources Management Association, 2011; Cristianini & Shawe-Taylor, 2000; Kulkarni, 2012), these require the user to either possess data that contains known classes or to make strong *a priori* assumptions about the number of groups before fitting any models. The few methods from the machine learning literature that do not require these assumptions to be made (e.g., hierarchical agglomerative clustering) still lack the sophistication to address issues related to observations not being independent (e.g., clustering user activity within users). These issues motivated our group's choice to use latent class analysis (LCA), and more specifically, multilevel LCA to address the lack of independence (Vermunt, 2003).

Using this approach, we are able to simultaneously classify each of the reports that a given user viewed (e.g., the within user class) and classify groups of users (e.g., the between user class). In other words, when we estimate the probability that a given educator belongs to a specific user group/type, we are also able to account for the type/classification of the interactions that user had with the system over the span of nearly a full academic year. Because these models are highly sophisticated mathematically, we refer interested readers to (Asparouhov & Muthén, 2008, 2014b, 2015; Henry & Muthén, 2010; Nylund, Asparouhov, & Muthén, 2007) for additional information on the mathematical derivation of these types of models as well as examples of their applications in various settings.

Given the limitations in most statistical software packages[1]—with regards to latent class modeling—we performed our data cleaning and preparation in Stata 14 MP8, used StatTransfer

[1] One limitation of the software selected is the number of distinct values a nominal scale measure can take. While we were able to reduce the number of report types to 13 distinct groups, Mplus only allows 10 discrete values to be used for a nominal scale variable. Ideally, we

13 to convert the Stata dataset to an Mplus dataset and input file template, and used Mplus 7.3 to fit the latent class models. Once the class membership was estimated, the data were reloaded in Stata to fit models to estimate the relationship(s) between data use aggregated at the school level and school level educational accountability outcomes. For additional information about this process, or to view the source code used during this process see Appendix C.

Model Building

Our strategy for building our model was to build from the most parsimonious models to more sophisticated models using a factorial approach where we varied: sample (all staff vs. users only), single vs. multilevel, whether covariates were included or excluded, and the number of latent classes a latent variable was allowed to take (two, three, four, or five)[2]. This led to fitting 32 distinct models from which we selected the best fitting model. To select the best fitting model, we used a combination of Akaike and Bayesian information criteria (AIC/BIC).

Because there were a significant number of non-users in the data set (n=3865), we tuned and tested the latent class analysis model by building models across the data with and without these observations included. By including records of non-users we could ensure that the observations would be correctly classified and could then test whether job role indicators

---

would want to model the report selection at a given point in time as a single multinomial variable, but due to this limitation we needed to create a saturated vector of indicators for each report type to serve as the within user dependent variables.

[2] In multilevel models the within user latent classes were fixed at five and only the between user latent classes were allowed to vary.

predicted class membership (e.g., would being a middle school classroom educator make someone more or less likely to be classified in a particular group?).

After testing single level models across the data with and without the non users, we then fitted more sophisticated models that allowed us to estimate relationships within and between users. No within user covariates were added to these models, but between user covariates (e.g., job type indicators or number of days visited) were added.

Results

When we consider the proportional amount of report types viewed by members of each of the user groups, there are clear differences in both the content (e.g., which report) and quantity (e.g., how much was the report viewed) in report type access (see Figure 1 for additional information). When we look at when the reports are being accessed (see Figure 2), we can also see that there are differences in the time dependency (e.g., one group is more likely to look at a report overall, and at different points in time one group could be more likely than another to view the report).

Model

Our model fit the data with a high degree of fidelity as summarized by the entropy statistic (0.964); a value of 1 would be a clear indication of over-fitting of the model to the data and values < 0.8 would typically be considered a poor fit to the data. Given the literature on LCA with covariates (Asparouhov & Muth, 2015; Asparouhov & Muthén, 2014a), we sought to first find an LCA solution without covariates which would be used to constrain the model parameters (e.g., the probabilities of selecting a given report conditional on being classified as a specific type of session) and tested the invariance of the classification based on education level and number of days the platform was used. Ideally, we would want non-significant relationships with these indicators since a significant relationship with a predictor would be an indicator of additional unmodeled error. We found a few instances where these covariates significantly predicted the between user groups. Both the total number of days the platform was used and the elementary school educator indicators were significant predictors of being classified in user group two. Conversely, the elementary school educator indicator was a significant predictor of not being classified in user group three (e.g., elementary school educators were significantly less likely to be classified in this group). Lastly, the total number of days the tool was used was also a significant predictor of being classified in user group four.

We included the marginal frequencies/probabilities of latent class membership for the within and between users groups in table 4. This table shows the total number of observations included in each class regardless of the class information from a class at a different level (e.g., the within user class probabilities do not factor the between user class probabilities). The conditional—or joint—probabilities (e.g., the probability of being classified as a given within user class for between user class 1) along with the other information about model fit, estimated parameters, and more are available in the additional resources listed in appendix C.

Relationship to Accountability Measures

After classifying users and their interactions with the data system, we also did some initial preliminary analyses of the relationship between the number of users in each of the between user classes and student level outcomes reported in the State of Kentucky's accountability system. To do this, we aggregated counts of users and counts of specific reports viewed by school. We then took both the 2014–15 indicators, as well as the first differences of those indicators from the previous year, and mined those data for possible relationships. Given the density and volume of the information, we wanted to provide an easier way to quickly evaluate and understand the relationships between the variables. Figure 2 is a heatmap of the correlation matrix. The red cells on the diagonal indicate the correlations between the variable and itself (always 1), and we used a divergent color palette from ColorBrewer (Brewer, 2015; Buchanan, 2015) to help highlight the differences between positive and negative correlations. Purple cells indicate a more positive correlation, while orange indicates a more negative correlation. Some of the more notable findings are the seeming lack of relationship that science points have with nearly all other variables in the correlation matrix and negative relationships with reading and math proficiency points and the number of users classified into different groups. The number of school staff classified in user group four is also interesting as it is nearly orthogonal to all status (proficiency) measures, but positively correlated with increases in proficiency from the prior year. Conversely the number of staff classified in user group five was positively correlated with status measures, slightly negative with change in secondary reading proficiency, slightly positively correlated with a change in math proficiency, and unrelated to all other changes in proficiency. We can also see that the percentage of ethnoracial minority students and students qualifying for free lunch is positively correlated only with the number of user group four, the same group that we observe having a positive relationship with changes in proficiency but not current proficiency.

There are also interesting patterns in the relationships between the number of times specific reports were accessed and the accountability system measures. In particular, reports five, six, seven, nine, and ten are unrelated to current proficiency measures (these are the larger collections of white space in the figure), but have a tendency to be positively correlated with changes in the proficiency points for those same variables.

In addition to estimating these correlations, we also applied data mining and machine learning techniques to the data to see if we could fit a linear model to the various accountability system outcomes. We regressed each of the reading and math proficiency outcomes on the set of permutations of user group indicators (Luchman & Cox, 2015), a Least Angles Regression (LARS), normalized/penalized regression methods (e.g., Lasso) (Mander, 2014), and best subsets regression (Lindsey & Sheather, 2010) methods to test these relationships. Although some of the results met the traditional threshold for statistical significance, we chose not to present the results here to avoid any possible confusion related to the interpretation of the results. In particular, given the small sample of schools (*n*=63) and the number of regressors, it was our opinion that the results were more likely to be spurious than true relationships between the variables. Instead, we advocate for future research directions that are more sensitive to our understanding of the underlying data production function (e.g., educators view data, modify instructional strategies for those children, children are assessed again, and the cycle restarts).

Lessons Learned and Future Directions

Lessons Learned. While this process requires significant time andresource investment at the start, we believe it is still a worthy endeavor that can better enable educational leadership to support classroom educators—as well as educators supporting each other. However, as techniques for user-segmentation in the context of web-application proliferate, it is likely that it will become easier and faster to conduct these types of analyses locally and move them into typical production systems used in the education sector.

Challenges. Getting the log file data could be the very first big challenge depending on how and where the data is stored. For example, data stored locally is usually easier to obtain compared to data stored through a third party vendor. Analysis of the data requires the technology stack to be integrated into the organization's existing IT infrastructure. Without support from IT staff, there is little hope to deploy these tools. However, the analysis provided by the tool can also yield valuable and actionable insight for IT staff.

Latent variable models are not only difficult to fit to the data (e.g., several models fit to our data during the model building process had fatal issues) but also challenging to discuss with audiences that may lack highly sophisticated understanding of statistical methods. However, these challenges also provide the necessary space to authentically engage the staff in your organization in the process of research and analysis. For example, if time had allowed we could have asked professional development coordinators or instructional coaches how they would label the classes as well as ask them about other behaviors that influenced their decision(s) about the labels.

One of the most potentially difficult challenges that could arise is potentially alienating segments of the staff if they feel their trust and/or privacy have been violated. To be clear, the goal with these types of analytical approaches is not holding end-users accountable for using the tool. Rather, it is a way that districts and/or states can hold themselves accountable for providing the necessary training and support to classroom educators, building leaders, and stakeholders to fully realize the greatest return on investment. Most importantly,it provides an empirical toolkit to more effectively reflect on our own understanding of how we use data and how we can improve our methods for using data to support children.

Solutions. Having a thorough understanding of what the end user is viewing (e.g., which report is accessed), when the user is viewing or using the system (e.g., when does the user login, access reports, and use interactive features), how the user navigates the system (e.g., which report is viewed first and length of time viewing each report), and the health of the system (e.g., amount of computing resources used to fulfill each report request, errors/warnings/failures, or transmission time) can provide IT professionals with the insight required to derive the greatest return on investment in their labor and infrastructure investments and maintenance. We have used these talking points to develop coalition around the analyses of these data with IT professionals and believe they can provide a helpful frame of reference with which to create a space for productive dialog.

Future Directions

Web analytic and data science. Several authors (Beasley, 2013; Berger & Fritz, 2015; T. Dinsmore & Chambers, 2014; T. W. Dinsmore & Chambers, 2014; Kaushik, 2009; Miller, 2014) provide robust coverage of machine learning, data science, and web analytic approaches to quantifying and analyzing the user experience with technology tools. Our ability to maximize the effective deployment and use of these tools rests on our ability to understand what does and doesn't work for all stakeholders as well as understanding how users typically use the tools. In particular, applying techniques such as Clickstream analytics (Kaushik, 2009) would provide us with a better understanding of how the user navigates through the system (e.g., where do they start or end, what do they do between then and how long do they stay at each of the intermediate points).

Miller (2014) also suggests investigating networks—or graphs—related to the platform/tool. In particular, studying the networks of professional learning communities and whether this is reflected in data use and analysis could provide data that is immediately actionable. For example, if a group of educators rarely is viewing data together or are viewing disparate reports, a district- or state-level intervention team could reach out and offer additional support and training to remove perceived barriers that may exist with regards to asking for help and assistance. Berger and Fritz (2015) are strong advocates of A/B testing. In other words, rather than simply making massive scale changes to these complex systems, we can randomly assign users to receive the same content via different interfaces and empirically test which interface is most preferable, easiest to use, and most likely to receive wider adoption.

Recommender Systems. While it is helpful to understand how users interact with your technology systems, one area that could provide significant benefits to your stakeholders is to build recommendation systems around your technology platforms. For example, we can analyze how the use of the data system—with regards to both the content (e.g., which report) and the sequence (e.g., the order in which the reports are viewed)—affects student performance and recommend users to view reports based on what would be most likely to have a positive effect on student learning outcomes. Developing systems like this can also be used to facilitate more robust integration of the data analysis and instructional staff through the use of mixed methods research designs that would directly integrate educator feedback into the recommender system.

Mixed Methods approaches. The quantitative analysis can show how users interact with your technology systems. However, it might be lack of capacity of predict how users want to use the technology systems. For example, the usage data indicates that only 25.42% of all staff in the district used the data platform at least one time, and of classroom educators only 47.96% used the data platform at least once. But does this mean only 25.42% of all staff in the district wants to use the data platform and of classroom educators only 47.96% used the data platform? In this case, surveys and/or qualitative methods can help answer the question. The qualitative results will provide information filling the gap of the actual usage and the intended usage. Based on the quantitative and qualitative feedbacks, you can then further investigate the reasons behinds the usage gaps.

What data can surveys and focus groups provide that log data cannot? The quantitative method is good at providing instant usage statistics patterns, but there is other critical information needed to make more informative decisions including technology access, training experience, and tech savviness. For example, in our case, district administrative need to know what are the reasons behind these patterns. Why is there only a very small percentage of people using the tool? Is it because there was alack of access, they did not know how to use it, or there was a lack of time to use it? Why did only certain groups of people use the tool? Is it because only they thought it is useful or because they had better training or access?

Why should you invest in surveys and focus groups to study this? While quantitative method can provide you a systematic analysis of the usage patterns, the qualitative method will provide necessary complementary information by digging into the reasons behind the patterns, users training experience, technology access, tech savviness, importance of the tools, and etc. Companies create instant survey applications, like BrightBytes's Clarity Survey, that make the qualitative survey relative easy to access, and theycan provide instant response to decision makers. With both quantitative and qualitative approaches, district decision makers will have more comprehensive information to make better decisions.

References

Asparouhov, T., & Muth, B. (2015). Auxiliary Variables in Mixture Modeling : Using the BCH Method in Mplus to Estimate a Distal Outcome Model and an Arbitrary Secondary Model, (21), 1–22.

Asparouhov, T., & Muthén, B. (2008). Multilevel Mixture Models. In G. R. Hancock & K. M. Samuelsen (Eds.), *Advances in latent variable mixture models* (pp. 27–52). Charlotte, NC: Information Age Publishing. Retrieved from http://www.loc.gov/catdir/toc/ecip0727/2007037392.html

Asparouhov, T., & Muthén, B. (2014a). Auxiliary Variables in Mixture Modeling: Three-Step Approaches Using M plus. *Structural Equation Modeling: A Multidisciplinary Journal*, *0*(June), 1–13. doi:10.1080/10705511.2014.915181

Asparouhov, T., & Muthén, B. O. (2014b). *Auxiliary Variables in Mixture Modeling: Three-Step Approaches Using Mplus* (No. 15). *Mplus Web Notes*. Los Angeles, CA. doi:10.1080/10705511.2014.915181

Asparouhov, T., & Muthén, B. O. (2015). *Auxiliary Variables in Mixture Modeling : Using the BCH Method in Mplus to Estimate a Distal Outcome Model and an Arbitrary Secondary Model* (No. 21). *Mplus Web Notes*. Los Angeles, CA.

Association, I. R. M. (2011). *Machine Learning: Concepts, Methodologies, Tools, and Applications*. Hershey, PA: IGL Global.

Baker, R. S. J. D., & Corbett, A. T. (2014). Assessment of Robust Learning with Educational Data Mining. *Research & Practice in Assessment*, *9* (Winter), 38–50. Retrieved from http://www.rpajournal.com/dev/wp-content/uploads/2014/10/A4.pdf

Baker, R. S. J. D., Corbett, A. T., & Gowda, S. M. (2013). Generalizing Automated Detection of the Robustness of Student Learning in an Intelligent Tutor for Genetics. *Journal of Educational Psychology*, *105*(4), 946–956.

Baker, R. S. J. D., & Yacef, K. (2009). The State of Educational Data Mining in 2009 : A Review and Future Visions. *Journal of Educational Data Mining*, *1*(1), 3–16. doi:http://doi.ieeecomputersociety.org/10.1109/ASE.2003.1240314

Beasley, M. (2013). *Practical Web Analytics for User Experience*. Waltham, MA: Morgan Kaufman.

Berger, P. D., & Fritz, M. (2015). *Improving the User Experience through Practical Data Analytics*. Waltham, MA: Morgan Kaufman.

Brewer, C. A. (2015). ColorBrewer2. Retrieved from http://www.colorbrewer2.org

Buchanan, W. R. (2015). BREWSCHEME: module for generating customized graph scheme files. Boston, MA: Statistical Software Components, Boston College Department of Economics.

Cristianini, N., & Shawe-Taylor, J. (2000). *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. New York City, NY: Cambridge University Press.

Datnow, A., Park, V., & Wohlstetter, P. (2007). *Achieving with Data: How high-performing school systems use data to improve instruction for elementary students*. Los Angeles, CA. Retrieved from http://www.newschools.org/viewpoints/AchievingWithData.pdf

Dinsmore, T., & Chambers, M. (2014). *Modern Analytics Methodologies: Driving Business Value with Analytics*. Upper Saddle River, NJ: Pearson FT Press.

Dinsmore, T. W., & Chambers, M. (2014). *Advanced Analytics Methodologies: Driving Business Value with Analytics*. Upper Saddle River, NJ: Pearson FT Press.

Gold, T., Lent, J., Cole, R., Kemple, J., Nathanson, L., & Brand, J. (2012). *Usage Patterns and Perceptions of the Achievement Reporting and Innovation System (ARIS)*. New York City, NY.

Hamilton, L., Halverson, R., Jackson, S. S., Mandinach, E., Supovitz, J. A., & Wayman, J. C. (2009). *Using student achievment data to support instructional decision making (NCEE 2009-4067)*. Washington, D. C. Retrieved from http://ies.ed.gov/ncee/wwc/publications/practiceguides/

Henry, K. L., & Muthén, B. (2010). Multilevel Latent Class Analysis: An Application of Adolescent Smoking Typologies With Individual and Contextual Predictors. *Structural Equation Modeling: A Multidisciplinary Journal*, *17*(2), 193–215. doi:10.1080/10705511003659342

Hung, J.-L., Hsu, Y.-C., & Rice, K. (2012). Integrating Data Mining in Program Evaluation of K–12 Online Education. *Educational Technology & Society*, *15*(3), 27–41.

Kaushik, A. (2009). *Web Analytics 2.0: The Art of Online Accountability and Science of Customer Centricity*. Indianapolis, IN: Wiley Publishing, Inc.

Kulkarni, S. (2012). *Machine Learning Algorithms for Problem Solving in Computational Applications*. Hershey, PA: IGL Global.

Lachat, M. A., & Smith, S. (2005). Practices That Support Data Use in Urban High Schools. *Journal of Education for Students Placed at Risk*, *10*(3), 333–349. doi:10.1207/s15327671espr1003_7

Lindsey, C., & Sheather, S. (2010). Variable Selection in Linear Regression. *The Stata Journal*, *10*(4), 650–669.

Little, T. D. (2013). *Longitudinal Structural Equation Modeling*. New York City, NY: Guilford Press.

Luchman, J. N., & Cox, N. J. (2015). TUPLES: Stata module for selecting all possible tuples from a list. Boston, MA: Statistical Software Components, Boston College Department of Economics.

Mander, A. (2014). lars. Boston, MA: Statistical Software Components, Boston College Department of Economics. Retrieved from http://fmwww.bc.edu/repec/bocode/l/lars

Means, B., Chen, E., DeBarger, A., & Padilla, C. (2011). *Teachers' Ability to Use Data to Inform Instruction: Challenges and Supports*. Washington, D. C. Retrieved from https://www.sri.com/work/publications/teachers-ability-use-data-inform-instruction-challenges-and-supports

Means, B., Padilla, C., & Gallagher, L. (2010). *Use of Education Data at the Local Level: From Accountability to Instructional Improvement*. Washington, D. C. Retrieved from https://www2.ed.gov/rschstat/eval/tech/use-of-education-data/use-of-education-data.pdf

Miller, T. W. (2014). *Web and Network Data Science: Modeling Techniques in Predictive Analytics*. Upper Saddle River, NJ: Pearson FT Press.

Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the Number of Classes in Latent Class Analysis and Growth Mixture Modeling: A Monte Carlo Simulation Study. *Structural Equation Modeling*, *14*(4), 535–569. doi:10.1080/10705510701575396

Ocumpaugh, J., Baker, R. S. J. D., & Rodrigo, M. M. T. (2015). *Baker Rodrigo Ocumpaugh Monitoring Protocol (BROMP) 2.0 Technical and Training Manual*. New York City, NY. Retrieved from http://www.columbia.edu/~rsb2162/BROMP.pdf

Panov, P., Soldatova, L. N., & Dzeroski, S. (2009). Towards an Ontology of Data Mining Investigations. In *Discovery Science* (pp. 257–271). Berlin: Springer.

Peña-Ayala, A. (2014). Educational data mining: A survey and a data mining-based analysis of recent works. *Expert Systems with Applications*, *41*, 1432–1462. doi:10.1016/j.eswa.2013.08.042

Romero, C., & Ventura, S. (2010). Educational Data Mining: A Review of the State of the Art. *Transactions on Systems, Man, and Cybernetics–Part C: Applications and Reviews*, *40*(6), 601–618. doi:10.1109/TSMCC.2010.2053532

Topol, B., Olson, J., Roeber, E., & Hennon, P. (2012). *Getting to Higher-Quality Assessments: Evaluating Costs, Benefits, and Investment Strategies*. Stanford, CA. Retrieved from http://edpolicy.stanford.edu/publications/pubs/747

Tyler, J. H., & McNamara, C. (2011). *An Examination of Teacher Use of the Data Dashboard Student Information System in Cincinnati Public Schools*. Washington, D. C.

Vermunt, J. K. (2003). Multilevel Latent Class Models. *Sociological Methodology*, *33*, 213–239. doi:10.1111/j.0081-1750.2003.t01-1-00131.x

Wayman, J. C., & Stringfield, S. (2006). Technology-Supported Involvement of Entire Faculties in Examination of Student Data for Instructional Improvement. *American Journal of Education*, *112*(4), 549–571. doi:10.1086/505059

Appendices

Appendix A. Descriptive Statistics

Table 1. Distribution of Job Types in FCPS

Job Type	Frequency	Percentage	Cumulative Percentage
Unknown	31	0.61	0.61
Accounting/Clerical/Operations	1178	23.29	23.90
Information Technology/Systems	65	1.28	25.18
District Administration/Central Office	238	4.70	29.89
School Administration	76	1.50	31.39
Special Education/Education Specialists	1251	24.73	56.12
Classroom Educators	2220	43.88	100

Table 2. Number of Classroom Educators By Grade Spans Taught

Grade Span	Frequency	Percentage
Unknown	55	0.01
Elementary	1148	22.69
Middle	506	10.00
High	511	10.10

Table 3. *Report Views by Job Type Classification*

	Unknown	Accounting, Clerical, & Ops	Information Tech/Systems	District Administration Central Office	School Administration	Special Ed Ed Specialists	Classroom Educators
ASG Class Report	80	0	11	193	299	201	2942
Class By RIT	38	0	16	147	54	159	2654
Class Report	245	0	99	447	171	808	7840
Class by Goal	8	0	0	54	62	57	1368
Class by Projected Proficiency	1	0	0	7	0	1	89
Des Cartes Query	8	0	0	89	7	26	569
District Summary	0	0	1	17	5	3	17
Grade Report	152	0	17	428	148	231	1065
MPG Student	0	0	0	0	0	0	16
MPG Sub-Skill Performance	0	0	0	2	0	0	4
MPG Teacher	0	0	3	5	0	1	41
PGID	2	0	0	9	1	9	419
Potential Duplicate Profiles	0	0	2	1	0	0	4
Profiles With Shared IDs	0	0	0	1	0	0	1
Projected Proficiency Summary	0	0	0	24	14	1	18
Student Goal Setting Worksheet	107	0	2	122	6	154	1906
Student Growth Summary	9	0	6	53	19	17	66
Student Progress Report	394	0	132	883	175	1930	8119
Students Without Reporting Attributes	1	0	0	1	0	4	2
Students Without Valid Test Results	11	0	75	98	4	115	215
Test Events By Status	1	0	0	4	0	6	7
User Roles	0	0	3	0	0	1	0
No Reports Viewed	0	1200	58	207	58	1156	1186

Table 4. Class Counts Based on Estimated Posterior Probabilities and Most Likely Latent Class Pattern

Estimate Type	Latent Class Variable	Latent Class Indicator	Frequency	Proportion
Estimated Posterior Probabilities	UGROUPS (Between Users)	1	5296	0.14686
		2	7978	0.22125
		3	7881	0.21856
		4	10230	0.28369
		5	4678	0.12964
	SESSION (Within Users)	1	3726	0.10333
		2	2297	0.06370
		3	8794	0.24387
		4	11633	0.32260
		5	9610	0.26650
Most Likely Latent Class	UGROUPS (Between Users)	1	5266	0.14603
		2	7895	0.21894
		3	7904	0.21919
		4	10268	0.28475
		5	4727	0.13109
	SESSION (Within Users)	1	3726	0.10333
		2	2297	0.06370
		3	8794	0.24387
		4	11633	0.32260
		5	9610	0.26650

Appendix B. Data Visualizations

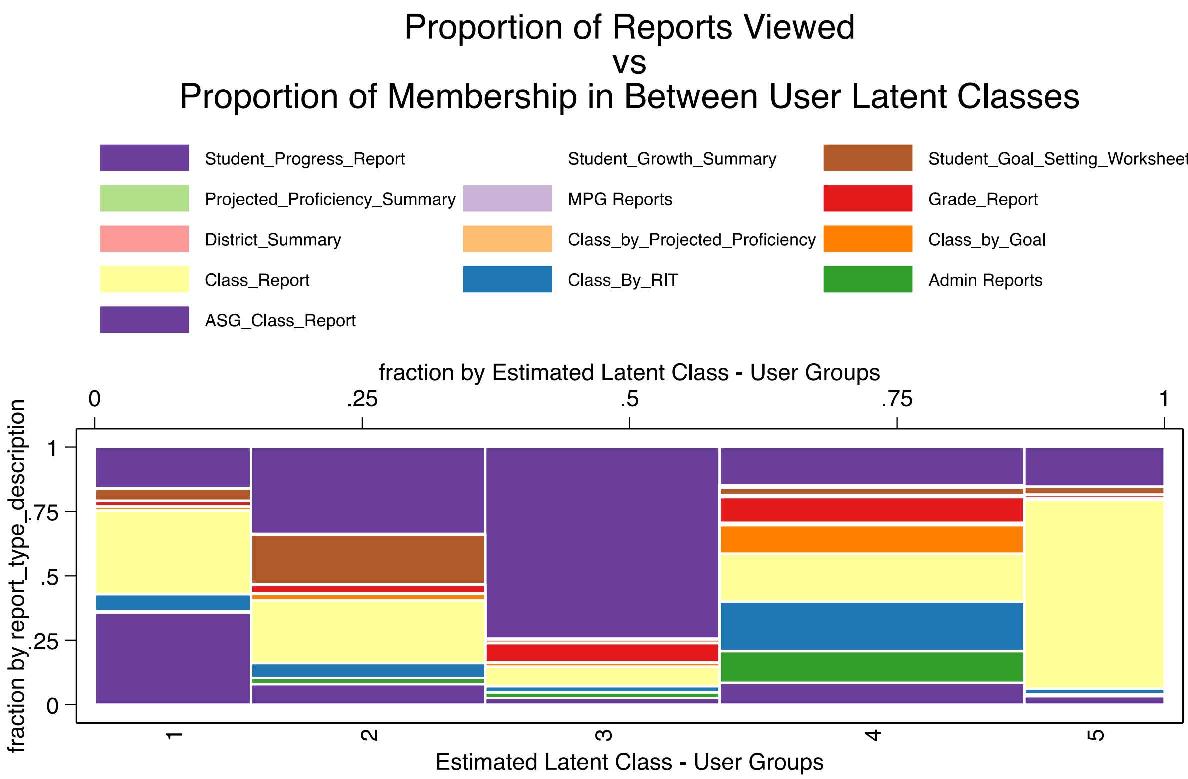

Figure 1. Proportions of Report Types Accessed by Estimated User Groups

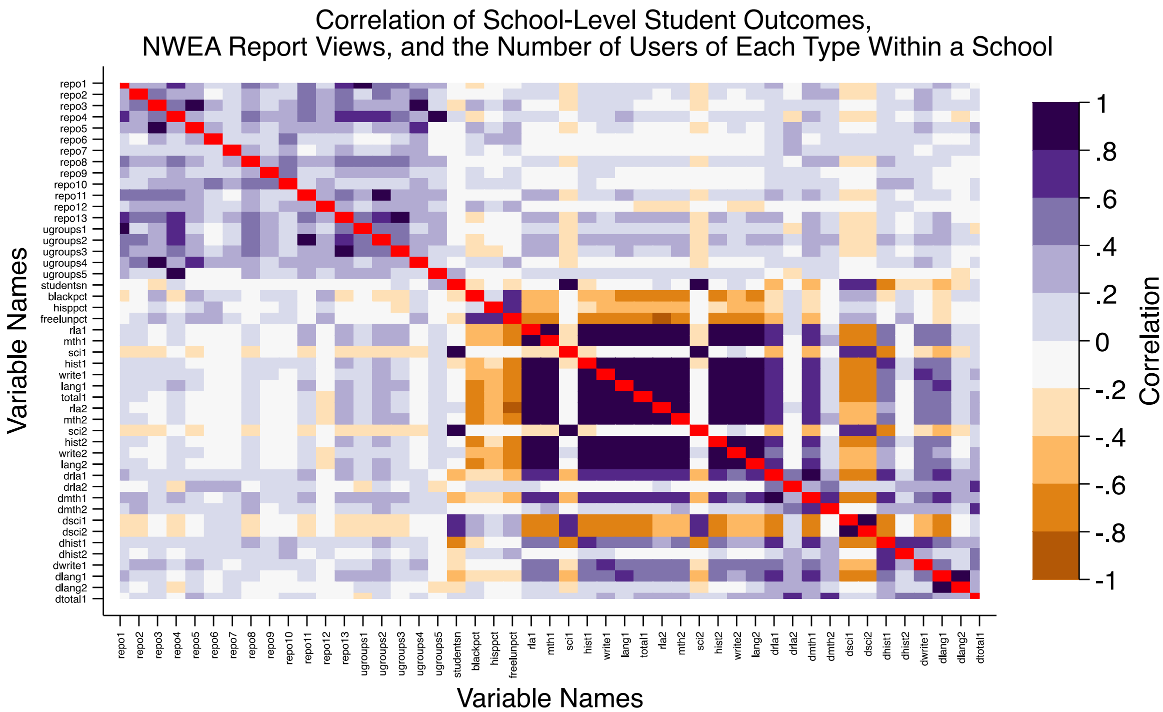

Figure 2. Correlations between report views, distribution of user group types, and school-level accountability measures

Stub	Meaning
Repo#	Report type (# = 1-13)
Ugroups#	Between user classification (# = 1-5)
Studentsn	Number of students in 2014–15 school year
Blackpct	% Black students in 2014–15 school year
Hisppct	% Hispanic students in 2014–15 school year
Freelunpct	% Free lunch eligible students in 2014–15 school year
Rla#[1]	Accountability points for Reading proficiency in 2014–15 school year
Mth#[1]	Accountability points for Math proficiency in 2014–15 school year
Sci#[1]	Accountability points for Science proficiency in 2014–15 school year
Hist#[1]	Accountability points for History proficiency in 2014–15 school year
Write#[1]	Accountability points for Writing proficiency in 2014–15 school year
Lang#[1]	Accountability points for Language Arts proficiency in 2014–15 school year
Total#[1]	Total accountability system proficiency points for 2014–15 school year
Drla#[1]	Change in accountability Reading proficiency points from 2013–14 school year
Dmth#[1]	Change in accountability Math proficiency points from 2013–14 school year
Dsci#[1]	Change in accountability Science proficiency points from 2013–14 school year
Dhist#[1]	Change in accountability History proficiency points from 2013–14 school year
Dwrite#[1]	Change in accountability Writing proficiency points from 2013–14 school year
Dlang#[1]	Change in accountability Language Arts proficiency points from 2013–14 school year
Dtotal#[1]	Change in accountability Total proficiency points from 2013–14 school year

[1] 1 = Primary level 2 = Secondary level

Appendix C. Tools for Others

To install the ELK Stack on your computer systems, we've created an installation script to help you get up and running a bit faster. Visit our installation tutorial and tools for instruction on how to use the script at: https://github.com/wbuchanan/elkStackInstaller.

To interact with some of our data and view our conference slides related to this report you can go to https://wbuchanan.github.io/capstoneProjectSDP; the underlying GitHub repository (https://github.com/wbuchanan/capstoneProjectSDP) also contains the source code we used to analyze the data in case you wanted to replicate and/or use our work as a starting point for your organization.