MLDS Center
Maryland Longitudinal Data System

Address 550 West Baltimore Street
Baltimore, MD 21201
Phone 410-706-2085
Email mlds.center@maryland.gov
Website www.MLDSCenter.org

Monthly Report for October 2015
Submitted November 6, 2015

Meetings

Center Staff met with the Division of Early Childhood Development within MSDE to resume discussions to understand and acquire early childhood data collected by that division. That data is needed to allow the Center to complete reporting required under the 2012 SLDS Grant (and Research Agenda item # 16) to *assess the training and retention of early childhood workforce in Maryland*.

At the September Governing Board meeting, the Center was asked to continue to explore the possibility of incorporating data from Maryland Department of Juvenile Services (DJS). Center staff met with staff from the Juvenile Services Education (JSE) Program at MSDE. The JSE program provides education services to youth detained and committed by DJS. JSE administers educational services in 14 DJS facilities statewide. JSE recently installed a new student information system through the 2014 SLDS grant award. JSE was very receptive to working with the Center. It was determined that the Center will bring together staff from JSE and DJS to further discuss the best way to provide outcomes on the youth served by DJS.

Center staff met with the Career and Technical Education office at MSDE in response to questions from stakeholders on the Youth Apprenticeship pilot currently in progress with MDSE and DLLR. The Center has begun to explore the possibility of including apprenticeship and youth apprenticeship data within the System.

Data Loads and Quality Assurance

The Center continues to encounter data quality issues which impacts the ability to accurately report data across sectors. Staff spent considerable resources in resolving and cleansing K12 identities in the Master Data Management database. Since K12 is usually the starting point for matching identities to other sectors, the majority of the past month was dedicated to these records. Staff further determined that identities that were unresolved from previously provided data can now be resolved. Center staff continue to work on these issues to ensure the data is as accurate as possible.

Specifically, this identity resolution included analysis of 1.5 million K-12 students from MSDE (academic years 2008-2014) attendance data to validate the identifiers for these students contained in the MDM database and resulted in:

- Creating 107,334 new system identities;
- Designating "likely valid" SSN's for 321,569 students using first/last initials from wage data;
- Merging 854 students with multiple SASID's who were determined to be the same person based on name and date of birth, race, gender, grade level and school enrollment;

- Performing cleanup in the MDM database tables to remove any incorrect names, dates of birth, SASID's and SSN's associated with these 1.5 million students.

Staff continues to load the remainder of the 2014-2015 academic year data as it is received from our partner agencies. Recent loads include 2014-2015 data sets from MSDE, scheduled quarterly wages from DLLR, and data files provided by MHEC.

Student Outcome Achievement Report

The Center received a request from MHEC for information in support of MHEC's preparation of the Student Outcome and Achievement Report (SOAR). The Center was asked to provide high school data for students graduating in the 2012-2013 academic year that are found in Maryland postsecondary institutions the following academic year (2013-2014). The aggregate data provided by the Center includes student math, English, and reading levels along with AP and IB exams for the purpose of determining how these data are associated with college-going and student outcomes.

Research

This month the Research Team has focused on exploring and cleaning up data in preparation for running analyses focused on four key research areas. The first area is the Dual Enrollment Report due to the Governor and General Assembly on December 15, 2015. The first step is to use the data to identify the dually (or concurrently) enrolled students. This means finding students who attending high school and enrolled in a postsecondary course. This has been more complicated than it might seem on the surface. For example, the data system includes postsecondary enrollments for the fall semester from MHEC but winter/spring enrollments it come from the National Student Clearinghouse (NSC). However, data provided through the NSC only includes high school seniors. The Research Team has developed a combined strategy that will accommodate the differences in data sources and will allow robust analyses for reporting.

Second, the Research Team is working on analyses of postsecondary degrees to workforce outcomes. Of particular interest is types of degree as it relates to average starting wages. This work is leading to some very useful findings. Further, this project has allowed Dan McNeish to train our new doctoral fellow, James Zheng, on the system using the postsecondary and wage data as a sample case. We continue to encounter problems where data within the System are not comparable with data publicly-reported by the agencies. Prior to reporting any numbers, a robust understanding of the reasons for the discrepancies is needed.

Third, the Research Team is working on analyses about developmental education (i.e., remedial course taking) across the state. Finally, the researchers are investigating high school exiters who do not attend postsecondary institutions, and their workforce outcomes.

2015 SLDS Grant

Dr. Woolley has applied for approval from the University of Maryland, Baltimore, Institutional Review Board (IRB) for the Synthetic Data project that is part of the 2015 State Longitudinal Data System Grant (SLDS). The first steps in that project will start once the IRB submission has been approved. In addition,

the paperwork and MOUs between MSDE, MLDSC, UMB and UMCP will need to be finalized before work can begin.

Research Series
The first Research Series of the 2015-2016 academic year was held on October 2, 2015. The presenter was Research Team member (and recently awarded Ph.D. recipient) Dr. Dan McNeish. Dr. McNeish presented on analytic strategies for nested or multilevel data that do not require multilevel analyses. The November 6th Research Series will cover the Synthetic Data project, explaining what synthetic data are, what they can do, why this is an exciting project, and all the benefits the project might bring. This Research Series will help the Center educate stakeholders about synthetic data and the project itself, and allow the Center to begin to receive input from stakeholders to ensure the synthetic data serves their needs.

System Development
Last month, it was reported that due to cost concerns, the Center was not going to utilize the disaster recovery and storage options or the firewall services that the Department of Information Technology (DoIT) provides through the Maryland Enterprise Solution. After the Center's cost concerns were reviewed by DoIT, DoIT and their vendor developed another proposal that is more in line with the Center's resources. Subsequently, the Center is reconsidering the DoIt solution and is continuing discussions with DoIT. .

Staff is also working on the following tasks:
1. Upgrading VNX (5500 to 5700) at the data center;
2. Scheduling the installation of the switches;
3. Upgrading of all components of Oracle (Oracle Database, Webcenter and OBIEE);
4. Finalizing the Data Security and Safeguarding Plan (DSSP) and hardening all the security passwords; and
5. Completing the migration from Oracle Warehouse Builder (OWB) to Oracle Data Integrator (ODI).

Website and Dashboards
The updated website design was launched along with a new dashboard on *Maryland Public High School Graduates - Initial Postsecondary Enrollments*. Additional tables and graphs will be added to the dashboard to include enrollments by county and institution type, as well as, race/ethnicity and gender breakdowns.

Staff is also developing prototype table structures and report narratives for addressing Research Agenda questions: 12-*What happens to students who start at community colleges and do not go on to 4-year institutions*, 18-*What are the workforce for Maryland students who earn a high school diploma but do not transition to postsecondary education or trainin*g, and 20-*What are the workforce outcomes of Maryland high school non-completers*?