Element	**Description**
Institution Name	Moraine Valley Community College Library
Institution Characteristics	Community college; Public FTE enrollment: 7,713 FTE library staff: ● 13.75 Librarians ● 11.5 Other staff ● 1.5 Student Employees
CARLI Counts Participant Name + Job Title	Marie Martino Systems & Catalog Services Librarian
Project Name/Title	Data Inventory and Collection Practices Analysis: A Pilot Project
Single Sentence Abstract	The Moraine Valley Library investigated the Circulation Department's data collection practices by way of staff surveys, interviews, and data repository mining and found various opportunities for improving the library's data-related systems and processes.
Motivation(s) for Project	Academic libraries are increasingly expected to use data to make improvements to resources and services, as well as illustrate their impact and value, both within the walls of their home institution and beyond. The Moraine Valley Library has been collecting, using, and reporting relevant data internally and to various external agencies through its many years of operation. However, a number of circumstances converged in a timely fashion to create an impetus for executing this data practices project. This included

	various staff members' renewed interest in assessment and data-driven decision making practices, as well as the recent implementation of institution-wide initiatives centered around data usage at the college, and, of course, the inception of the CARLI Counts program. Furthermore, it has been at least 10 years or more since the library has conducted any sort of data audit or formal assessment of the processes and policies related to the varied components of the library's data life cycle. The Circulation Department, an area that generates a considerable proportion of the library's data, seemed like an appropriate place to pilot a formal evaluation of the library's data collection practices. Additionally, there was some evidence that legacy processes for generating certain data were being utilized and accompanying questions about their long-term viability had surfaced. As a result, a key component of this project included a systems process review. While the overall goal of what will be an ongoing, library-wide project is to build a better structure that allows the library to use data more effectively and more easily construct an engaging narrative about our library, the immediate objectives of this pilot project were to: ● Begin the creation and maintenance of a current data inventory guide ● Find data collection gaps and eliminate any data duplication ● Improve data collection workflows/processes Lastly, this pilot project will be evaluated and adjusted as needed to be reproduced in other areas of the library.
Partners and Stakeholders	Circulation Department staff members Dean, Learning Resource Center

	Librarians and other library staff
Inquiry Question	How will conducting a structured assessment of the library's data collection practices positively impact the library's ability to use data more effectively and allow us to better tell our story?
Study Participants/Population	Circulation Department Staff Dean of Library
Method(s) of Data Collection and Analysis	The project coordinator (i.e. the CARLI Counts participant listed above) used various methods to collect information for this project over the course of 11 months: Literature Review--The literature provided context for the project and models of similar projects completed at other academic institutions. Additionally, a couple of generous colleagues from other CARLI Counts institutions shared internal documents that served as great examples for building the data inventory guide. Survey--This tool was used as a means for identifying library staff members who collect and/or maintain data. It established a criteria for who on staff should be interviewed at length about data collection practices. It was distributed to all staff electronically via Google forms. For the purposes of this pilot project, only Circulation Department staff were marked for follow-up. (Note: Non-Circulation Department staff answers are being saved for when the remainder of the project it set to move forward.) The survey also functioned, in this way, as a means to raise awareness about the data project underway. In-person, semi-structured interviews--Next, staff members who were flagged from the initial survey participated in an interview process. The project coordinator scheduled 30 minute or 1 hour meetings with participants, depending on the amount of data

	each had claimed responsibility for in their survey answers. Interviews took place at the interviewee's workstation. The interview instrument was a semi-structured tool with 5 main questions. The first question, which asked interviewees to describe the data they collected/maintained, had a number of sub-questions such as how frequently it was collected, who it was reported to, what format was it saved in, what was its purpose/value, etc. Another question asked the staff member to identify all of the software systems they used to generate data. Another question asked them about any problems they have had with the data or the collection process involved. The next question prompted staff to identify any gaps in the data the department collects. Lastly, they were asked if they had anything else they wanted to share about their experiences related to data collection practices. System workflow/process analysis--This was the final phase of the interview, where applicable, where library staff were asked if they would be willing to demonstrate their system-generated data processes to the project coordinator, whose permanent role is the systems librarian. Each participant agreed to participate in this activity. The project coordinator observed, took notes, and where needed, flagged less efficient system processes for further investigation, especially if it was clearly deemed a "legacy" process. Collected sample data--Each participant provided the project coordinator with samples of the data they collected, post-interview. These were a mix of the raw data outputs they received from the system and/or spreadsheets that were created by staff themselves. Document mining--The project coordinator mined known data repositories and file storage locations. This made it possible to compare various data points with what was collected from the participants and

	flesh out any missing data sources for the data inventory guide. Based on the information gathered, two main deliverables could be produced: ● A current data inventory guide for staff ● A report that outlines the results of the investigation
Findings	First, as staff at all levels began to have conversations around this project and this concept of "data collection," it became clear that staff members had different ideas about what this meant with regard to their duties. In the immediate sense, a working definition was needed, but, more broadly, this may have implications worth exploring as the library works to become more of a data literate entity. Within the Circulation Department, there are 14 staff members, all of which who work with data in some capacity. However, out of that number, 6 staff are responsible for maintaining specific data and reporting it in regular intervals to the Dean. The data inventory guide now serves as a map to which staff member collects what data point and in what format it lives, where that data is saved/stored, and how it is used. It was learned that saved data exists in a mix of print and/or digital formats, depending on what it is and who is responsible for maintaining it. Additionally, the Dean and another professional staff member outside of the Circulation Department are generally responsible for more substantial analysis and reporting out circulation (and other data) both internally, college-wide, or to external agencies such as ACRL, ICCB, HLC, NCES, etc. No duplicate data collection was found among the circulation staff, but staff members did share some perspectives on where there may be gaps. This

	information will be addressed and shared in the final report, with the hope that the results inspire departmental discussions. A few system inefficiencies and/or questions related to generating system statistics were revealed. An example of this, in the immediate sense, pertains to how a certain set of statistics for reserve materials are generated and subsequently used and distributed to faculty. The process for getting that data from the integrated library system has long been a time-consuming effort due to the limits of the system for a number of years. But updates to the system within the last couple years now allow for an easy and efficient way of accessing and sharing that data. The process has since been completely revamped, shaping what was a week-long project into one that can be completed in a few hours. With further regard to longer term system goals, an investigation into fully automating some statistical gathering processes would create even more efficiencies and help free up staff time. For Circulation data alone, there are at least 4 “central” locations data where is stored, not including individual staff workstations. Because much constructive information was already gleaned from this pilot investigation, it is clear that a library-wide data collection analysis project would prove worthwhile and beneficial, so this project will be expanded into other areas.
Use of Findings	The findings of this project will be used to implement system and data collection process improvements and there are a number of sub-projects being planned. The project will continue and expand into other library departments.

	The final report will eventually be used to generate important discussions about data. First, these discussions can raise awareness about the use of data within the organization and illustrate each staff member's relationship to the data they collect. Additionally, within the different departments, specific data points can be assessed. What is missing? What might be removed? It may also be helpful for each area to articulate clear policies related to its handling of data through each phase of its life cycle and come up with an ongoing schedule for performing data audits. As this project has already revealed that data can be siloed within different departments, a plan for centralizing data storage is on the docket. Once data storage is centralized, the next step is conducting an investigation of open source software systems for data summarization and visualization, as these are helpful tools for further integrating data usage to drive decisions.
Next Steps and Other Results	As explained above, the library will continue to expand this project into other areas beyond the Circulation Department. In the big picture sense, it is clear that the library has more work to do to better incorporate data usage into its culture. As Perry and Petersen state in their 2017 article: *Providing an inventory and access to data is only the first step . . . Developing a culture that sees data analysis as a foundational element requires champions willing to demonstrate skills and help to train fellow staff members* (576). While, the library is definitely more mindful about how it uses data to drive decisions and show impact, this project is also a single, active step in a series of the many steps necessary to build a culture that embraces data usage in a more holistic and productive way. This will involve dedicated effort and

	an investment in the resources and training necessary to support such a culture.
Additional Reflections	The methodology used in this project was effective, albeit very time-consuming, especially running the interviews, which lasted anywhere from 15 minutes to 2 hours depending on the data collection practices being discussed and observed. The initial surveys can be used to identify staff who might require longer interview sessions. Prepare interviewees for this time commitment at the time meetings are being scheduled. It is also helpful to make these considerations as the timeline is being built to avoid time-creep. That said, this model may work well for a smaller library, but would very likely have to be modified for a larger organization or one that is experiencing low staffing issues. Also, good communication can help set clear expectations for interviewees. Sending a quick email or having a short conversation with the staff being interviewed beforehand helped prepare them for the questions being asked, and thus, set the stage for more productive conversations. Also, proactively addressing the purpose of asking staff to demonstrate their data collection processes made them less resistant to it. Rather than it being perceived as an assessment of their methodology or their skillset, it was viewed, accurately, as an audit of the system process.
Timeline	February - March 2019: Literature review April 2019: Distribute survey, analyze survey May - November 2019: Schedule and run interviews December 2019 - January 2020: Data repository mining, compiling deliverables (Phase 1 of data inventory guide, write Circulation Department section of final report)

Bibliography/Works Cited	Matthews, J.R. (2018). The evaluation and measurement of library services. Santa Barbara, CA: Libraries unlimited. Perry, M., & Petersen, G. (2017). Do we collect that information and if so, how can I access it? Designing a statistics depository. In Baughman, S. Hiller, S., Monroe, K., & Pappalardo, A. (Eds.), *Proceedings of the 2016 Library Assessment Conference (*pp. 574-577). Washington, DC: Association of Research Libraries. Retrieved from https://www.libraryassessment.org/past-conferences/2016-library-assessment-conference/2016-proceedings/
Appendices	As needed. Might be visualizations of the data, survey instruments, etc.

Note: Report will be submitted as a single PDF.