

Cultivating Collaborative Action: National Supply and Demand Data Sharing

Michael Spiess

Professor Emeritus
California State University, Chico
College of Agriculture
Chico, CA 95929-0310
(530) 521-6866
mspiess@csuchico.edu

Amy R. Smith

Associate Professor and Program Leader
University of Minnesota
146 Ruttan Hall
St. Paul, MN 55108
(612) 624-6590
arsmith@umn.edu

Daniel Foster

Associate Professor
The Pennsylvania State University
212 Ferguson Building
University Park, PA 16802
(814) 863-0192
foster@psu.edu

Rebecca Lawver

Department Head & Professor
Applied Sciences, Technology and Education
Utah State University
1498 N 800 E | ASTE 101C
UMC 2300
Logan, UT 84321
(435) 797-1254
rebecca.lawver@usu.edu

Cultivating Collaborative Action: National Supply and Demand Data Sharing

Introduction

As evidenced by the American Association for Agricultural Education (AAAE) Research Agenda (Stripling & Ricketts, 2016), a collective and critical issue facing agriculture is access to a sufficient and qualified workforce. Having a supply of qualified teachers is, and has been, a well-documented and pressing concern for school-based agricultural education (SBAE). As such, since 1965, AAAE has collected supply of and demand data for SBAE teachers across the U.S.

Since 2014, National Ag Ed Supply and Demand (NSD) data collection has been facilitated by a team of AAAE members. Annually, an executive summary has been provided to AAAE membership and published for broader use, with three-year reports available at the conclusion of each three-year cycle. While presenting discipline-related data through periodic reports to the profession is commonly accepted practice, a gap exists between current dissemination practices and what is possible leveraging modern digital technology.

Structural capacity limitations and human resource availability have, in some cases, prevented timely responses to data requests from AAAE members and stakeholders. As such, the team was compelled to explore open-access possibilities and consider open research efforts. Open access refers to immediate and online availability of research output, like annual reports, and data with rights to use outputs in a digital environment for research (Springer, 2023). Open access data has accelerated the pace of discovery with enhanced interdisciplinary conversations and collaboration, increased citation and usage, and fostered broader public engagement.

Government agencies commonly publish datasets on the internet, but it is less common for a discipline to publish research data in this way. While government data may be useful, it is often not in a format conducive to further analysis and/or may be poorly documented. Relational database systems (RDBS) were first described in a seminal article by Cobb (1970), who posited that all data could be defined or represented as a series of relations with or to other data. A spreadsheet by itself is not a database (Dilling, 2020); a database enables creation of a multidimensional structure to cleanly and accurately contain data. Presenting NSD data in a way that allows for new research techniques and methods could enable both faculty and student scholars in AAAE to engage in temporal analysis. It may also allow for the identification of more actionable research as we seek an understanding of and solutions to the long-standing challenge of ensuring a sufficient and qualified workforce of SBAE teachers.

How it Works

From 2014 to 2022, data was collected using two separate Qualtrics instruments sent to different populations. A supply instrument, collecting information about agricultural education program completers, was sent to agricultural teacher education program representatives, while the demand instrument was sent to agricultural education state staff representatives. Qualtrics data output was stored in Excel spreadsheets. To aggregate data, existing data had to be standardized, as there was slight variation between years. Once standardized, the data was imported into a Microsoft Access (MS-Access) relational database as two tables (supply and demand). MS-Access was chosen since it is a common desktop application and is easily integrated into the web application. Once imported, a table was created for regional grouping of data (e.g., AAAE, National Association of Agricultural Educators (NAAE), and FFA regions). Additional tables drive the web application for chart generation and report lists. A web application was written in ASP.NET to retrieve the data from the database. Google Chart, a free

data visualization tool, was used to drive the visual representation of the data in chart and map formats. The application is a data-driven interface which minimizes the need for maintenance.

Using simple Structured Query Language (SQL) queries, data is extracted and presented in chart or tabular form. Using SQL, data can be extracted from either table or combined for more complex analysis. In addition, common metrics can be calculated. Since data extraction is driven by queries, adding, or editing a report is a simple task requiring only a few minutes. The nature of the RDBS makes connecting the two datasets simple. Tabular data can be viewed on a web page or downloaded in an Excel readable (CSV) file. The user may choose from a variety of queries in a drop-down list and specify a year or state. The website also provides a data dictionary (stored in the database) and PDF copies of instruments used. The web application is driven by the data in the database so updates do not require any additional coding. To update either queries or the data, the MS-Access database is edited and then uploaded to the website.

Results To Date

The web application is available at <http://aaae.agedweb.org/nsd>. Since development, the data has been updated twice, and a total of 22 tabular queries and 14 chart queries have been developed. The process of combining annual data, from 2014 to present, into a common format and database enables more convenient access to longitudinal research, allowing for quick extraction. Extracted data can be analyzed using common research tools like SPSS, SAS, and Excel. A brief manual was written describing the application and how to update the database. A “user’s guide” is also provided to aid researchers in analysis of downloaded data. Data is available in much more detail than written reports for additional analysis.

A challenge to the design was dealing with incomplete data. Flags were added to the data to indicate valid data. These flags are used by the queries to select only valid records. The ease of extraction has also been helpful in connecting the dataset to other sources of data such as FAEIS, NCES, and others. Making NSD data more accessible is an initial step toward encouraging collaboration and parallel lines of inquiry in our discipline for heightened research impact.

Future Plans/Advice to Others

The immediate plan is to update the MS-Access database as data is collected. As such, consistent terminology, variable naming, and spreadsheet structure is important. For similar projects, this process may be helpful to organize data sets allowing for additional research. SQL queries should be developed around the needs of potential users; requests from AA AE members and other Ag Ed stakeholders may encourage the development of queries. Selection criteria might commonly be on date and other appropriate groupings. With consideration of privacy, data should be aggregated as necessary to hide information not in the public domain.

Costs/Resources Needed

Costs for this innovation include data preparation, building the web app, hosting, and maintenance. The cost of data preparation will be highly variable depending on data format and uniformity. For this project, approximately 16 hours was required to assemble all years into the database. The web application was built in about 8 hours; due to the technical nature of this task, it may need to be contracted. A total of 4 hours were spent on documentation and outlining the process for updates. Hosting services commonly cost \$200/year if a campus host is not available. Ongoing maintenance is required if/when additional data is collected. The time for updates is estimated as 1-2 hours/year (dataset).

References

- Codd, E. R. (1970). A relational model of data for large shared data banks. *Communications of the ACM*.13:377–387.
- Dilling, T. J. (2020). Artificial Intelligence Research: The Utility and Design of a Relational Database System. *Advances in radiation oncology*, 5(6), 1280–1285.
<https://doi.org/10.1016/j.adro.2020.06.027>
- Lawver, R. G., Foster, D. D., & Smith, A. R. (2018). *Status of the U.S. Supply and Demand for Teachers of Agricultural Education, 2014 - 2016*. <http://aaaeonline.org/Teacher-Supply-and-Demand>
- Springer Nature. (2023). *The fundamentals of open access and open research*.
<https://www.springernature.com/gp/open-research/about/the-fundamentals-of-open-access-and-open-research>
- Stripling, C. T., & Ricketts, J. C. (2016) Research Priority 3: Sufficient Scientific and Professional Workforce That Addresses the Challenges of the 21st Century. In T. G Roberts, T. G., A. Harder, M. T. & Brashears, (Eds). *American Association for Agricultural Education national research agenda: 2016-2020*. (pp. 30-35). Gainesville, FL: Department of Agricultural Education and Communication. Retrieved from:
http://aaaeonline.org/resources/Documents/AAAE_National_Research_Agenda_2016-2020.pdf