Cultivating Collaborative Action: National Supply and Demand Data Sharing

Michael Spiess[1]
Amy Smith[2]
Daniel Foster[3]
Rebeca Lawver[4]

Introduction

As evidenced by the most recent edition of the American Association for Agricultural Education Research Agenda (Stripling & Ricketts, 2016), a collective and critical issue facing agriculture and agricultural efforts is access to a sufficient and qualified workforce. Additionally, having a supply of qualified individuals is, and has been, a well-documented and pressing concern for all of education, in particular, for school-based agricultural education (SBAE). As such, since 1965, the American Association for Agricultural Education (AAAE) has collected data regarding the supply of and demand for SBAE teachers across the U.S. From 2014 to present, National Ag Ed Supply and Demand data collection efforts have been led by AAAE members Rebecca Lawver, Utah State University; Amy Smith, University of Minnesota; and Daniel Foster, The Pennsylvania State University. Since 2014, an annual executive summary of data collected has been provided to AAAE membership and published for broader use, in addition to three-year reports upon the conclusion of each three year cycle. While presenting discipline-related data such as this, with periodic reports back to the profession, is a commonly accepted practice, a gap exists between the 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 for 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 the immediate and online availability of research output, like annual reports, and data with rights to use outputs in a digital environment for open research efforts (Springer, 2023). Open data, in other aspects of scholarship, has accelerated the pace of discovery with wider collaboration, increased citation and usage, broader public engagement, and increased interdisciplinary conversations.

Government agencies commonly publish datasets on the internet, but it is less common for a discipline to publish research data for others to access. While government data may be useful, it is often not in a format conducive to further analysis and/or may be poorly documented. With the rapid expansion of artificial intelligence utilization (see ChatGPT conversations in mass media), the possibilities of utilizing a relational database system (RDBS) for research purposes are exciting. RDBS was first described in a seminal article by Cobb (1970), positing 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 these data. Presenting National Ag Ed Supply and Demand data in a way that provides opportunity for utilizing new research techniques and methods could enable both faculty and student scholars in AAAE to engage in temporal analysis. Additionally, it may allow for the identification of more actionable research by practitioners as we seek an understanding of and solutions to the long standing challenge of available SBAE teachers.

This poster highlights an opportunity for collaborative action to create actionable data for policy implementation regarding the supply of SBAE teachers. While an effort to make National Ag Ed Supply and Demand data more accessible, this could be an initial step toward applying open data sources to other lines of inquiry in our discipline for heightened research impact.

How It Works

Challenges

The survey data contains a number of sections (e.g.; Race) some of which may be incomplete. To avoid using incomplete data sections were flagged to omit incomplete data when queried.
Gathering the initial data from individual annual spreadsheets required developing standard field names (headings).

Using the Data

Data may be downloaded into Excel or a statistical package for analysis. Selections are provided by year and state. Extracted data can be analyzed using common research tools like SPSS, SAS, and Excel. A ”user’s guide” is also provided to aid researchers in analysis of downloaded data.

Relational Database

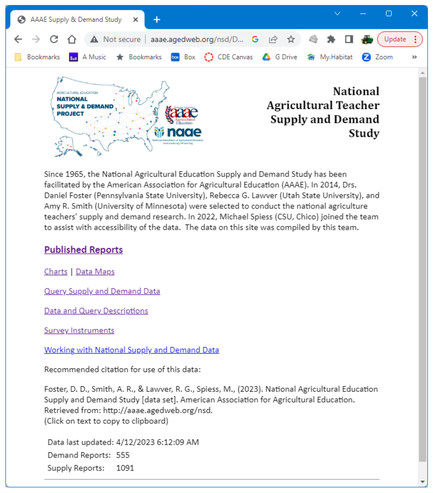
The use of a relational database has large advantages over using Excel or other “flat file” systems. Supply and Demand data are easily connected (“joined”). Summarization by state, region, or nationally is easily accomplished. Data can also be easily connected to external data using common data elements.

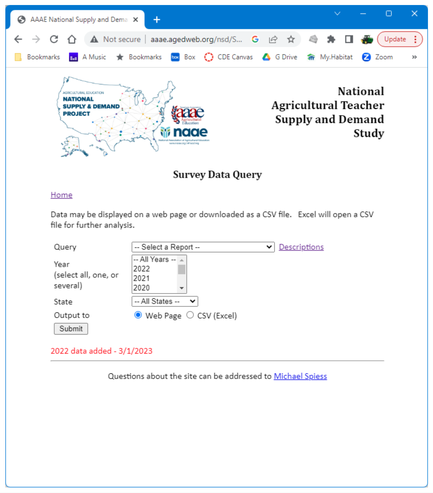
Microsoft Access (MS-Access) was chosen because it is a common desktop application and easily interfaced to the web application.

Data is extracted from the database using Structured Query Language (SQL). MS-Access provides a graphical interface to create queries which reduces the need to write in SQL.

Data Driven “App”

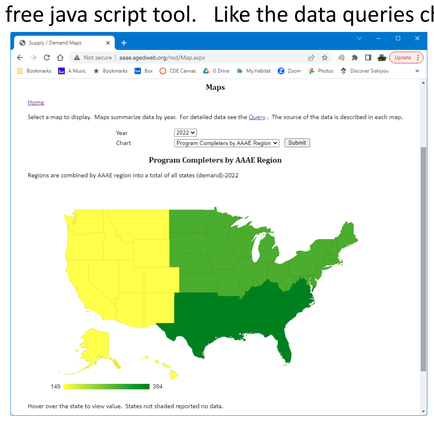
The web application was built using ASP.NET. The web-based app is entirely data driven. Using an Access database to house the data allows maintenance of the site to be accomplished using the Access desktop application. Data as well as the queries, charts, and maps are all stored in the database. Adding data is simply an import of new survey results into the database.

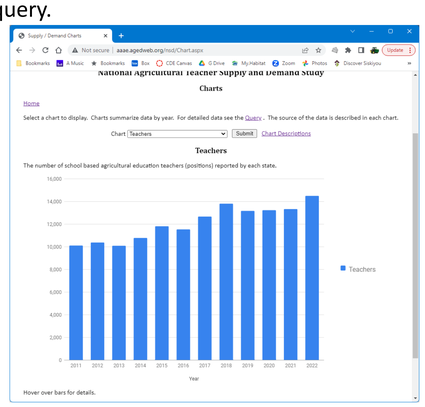




Visualization

To introduce a user to the data sample charts and maps are provided. These are rendered with Google Charts; a free java script tool. Like the data queries charts are driven by a SQL query.





Results To Date

Since development, the data has been updated twice, and a total of 33 tabular queries and 23 chart and 11 map 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 by allowing for quick extraction. Data is available in much more detail than written reports and data is available for additional analysis. A brief manual was written describing the application and how to update the database. Data requests are now handled by referring inquires to the web site.
The ease of extraction has also been helpful in connecting the dataset to other sources of data such as FAEIS, NCES, and others. Making the 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

The immediate plan is to continue to update the MS-Access database as data is collected. The system will be used for future reporting for the project to insure standardized data presentation. For similar projects, this process may be helpful to organize data sets allowing for additional research. SQL queries should be developed around the likely needs of the potential user; in this case, requests from AAAE members and other Ag Ed stakeholders may encourage the development of specific queries. Selection criteria might commonly be on date and other appropriate groupings. With consideration of privacy issues, data should be aggregated as necessary to hide information not in the public domain.

Cost

Costs for this innovation can be segmented into data preparation, building the web app, hosting, and ongoing maintenance. The cost of initial data preparation will be highly variable depending on the format and uniformity of the data. For this project, approximately 16 hours was required to assemble all the years into the MS-Access database. The web application was built in about 8 hours; due to the nature of this more technical task, it may need to be contracted. A total of 4 hours were spent on documentation of the application and outlining the process for updating the database. Hosting services would commonly cost $200/year if a campus host is not available. Ongoing maintenance is only required if/when additional data is collected. The time to update the database 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

[1]California State University, Chico, [2]University of Minnesota, [3]Pennsylvania State University, [4]Utah State University