Original Research

Empowering Those Who Seek to End Hunger through Collaboration and Innovation

Jen-Li Ko[1], Jeff Joslyn[2], Ganhua Lu[3], Jason Palmer[4], Malcolm Charles[5], Aishwarya Sanganalu Mattha[6], Bob Parsons[7], Ron Tatum[8], Carey Redmann[9], and Walter Bialkowski[10]

[1]Department of Languages, Literatures, and Cultures, Marquette University, [2]Feeding America Eastern Wisconsin, [3] Graduate Student, Marquette University, [4]Graduate Student, Marquette University, [5]Graduate Student, Marquette University, [6]Graduate Student, Marquette University, [7]Feeding America Eastern Wisconsin, [8]Feeding America Eastern Wisconsin, [9]Business Development Manager, Feeding America Eastern Wisconsin, [10]mailto:Department of Computer Science, Marquette University

Editor: Valerie L. Holton, Ph.D.

Abstract

Feeding America Eastern Wisconsin (FAEW) distributed 84% more food to community members in need during the COVID-19 pandemic than in the prior year. Though systems were in place to manage food receipt and distribution data, social distancing requirements and technological barriers revealed inefficiencies in utilizing this data. In pursuit of data-driven decision-making and in the context of a global pandemic, FAEW partnered with Marquette University data scientists through an industry-supported grant. Applying newly learned skills in Business Intelligence, students produced detailed reports of data cleanliness in FAEW's source systems to improve underlying data quality and better support analytic efforts. Additionally, students synchronized human-centered design thinking and visual analytics to produce an interactive application to optimize inventory management, storage availability, and product distribution. Finally, students are utilizing business analytics techniques such as supervised and unsupervised data mining to provide new insights about food receipt and distribution patterns that will have a sustainable impact on FAEW operations. This unique partnership offers students experiential

learning opportunities, tangible data science solutions that FAEW will use to ensure best practices, and real-world solutions to collaboratively end hunger in our communities.

Keywords: COVID-19, Data Science, Business Intelligence, Data Analytics, Food Distribution, Community engagement

Introduction

Thirteen million U.S. households face food insecurity (>10% of all households nationwide), representing uncertainty in having, or inability to acquire, enough food to meet the needs of household members due to insufficient money or other resources (U. S. Department of Agriculture, 2021). Feeding America Eastern Wisconsin (FAEW) served more than 400,000 people in 35 Wisconsin counties (10.2% of the population) before January 2020; however, in the subsequent one month period, the U. S. Bureau of Labor Statistics (2021) reported an increase of 10,000 unemployed people (approximately quadrupling the number of unemployed). This dramatic spike in unemployment exaggerated underlying disparities in our community, forcing many additional families to rely on the public service FAEW provides to receive healthy and nutritious food. Between February and April of 2020, FAEW saw a 67% increase in food distribution compared to the same timeframe in 2019. Much of FAEW's food is donated by wholesalers, retailers, and manufacturers, leaving most of the donated food received near the end of its shelf-life. Effectively managing this increase indefinitely required innovation and resources that FAEW did not have.

Marquette University's (2021) prestigious President's Challenge award supports innovative, interdisciplinary, and collaborative work that addresses critical issues and opportunities facing the City of Milwaukee. Marquette University faculty and students formed a unique alliance with FAEW leadership that brings together the field of data science and the mission to end hunger. Through experiential learning, faculty-mentored students developed the business intelligence, business analytics, and visual analytics tools FAEW needs to end hunger more effectively in our communities, which is illustrated in Table 1. Students (a) optimized current visual analytics practices using licensed software, (b) performed high-level data interrogation on legacy information within FAEW's database, (c) created an entirely de novo interactive application to support business decisions, (d) attempted to utilize historical information to identify determinants of food purchases, and (e) transformed a tabular hunger needs formula into an interactive application to support ongoing funding efforts. This complex project empowered students to utilize cutting-edge data science techniques to address meaningful and contemporary real-world challenges in an environment where the highest levels of corporate leadership are highly engaged and directly impacted by the accordant results.

Table 1. A list of data science initiatives supported through this collaborative, interdisciplinary, and innovative partnership. Each project was led by one student or groups of students based on their interests.

Challenge	Data Science Solution	Description
Optimizing Dashboard Functionality	Business Intelligence	Elevation of the level of report automation to save time spent creating Tableau dashboards.
Legacy Data Cleaning	Business Analytics	The inspection and correction of inaccurate or missing values within a dataset to improve data quality, and therefore the overall productivity for analytic purposes.
Interaction with Company-Wide Data	Visual Analytics	Interactive web application designed to give users the ability to glean meaningful insights in previously specified areas of interest within FAEW.
Purchased Product Optimization	Business Analytics	Supervised and unsupervised data mining designed to reveal patterns in the data, leading to the optimization of purchased products.
Hunger Needs Formula	Visual Analytics	Interactive application giving users the ability to manipulate variable values to estimate Total Ask and Total Estimated Pounds.

Novel Experiential Learning Pipeline

Learning data science techniques requires data, and often, students seeking training in data science have data sets available to them through work, research, or other avenues. Some students, however, wish to gain access to new data sets that support their learning. Two leading faculty members and one student had a first formal meeting with leaders from FAEW and toured the warehouse at FAEW shortly before the COVID-19 pandemic. To end hunger, especially during the COVID-19 pandemic, faculty and leaders at FAEW proposed this collaboration project. In this case, students were invited to express their interest in participating in a research team seeking to aid FAEW. Students selected for participation in the program could concisely articulate their interests and qualifications. Selected students (n = 5) executed non-disclosure agreements with FAEW, and created a secure data portal to exchange data between Marquette University and FAEW. In addition to receiving hourly pay, students could submit qualifying work for course credit in Marquette University's Data Science curriculum with instructor approval. Students, key leaders at FAEW, and faculty advisors met monthly to review the

progression of student work, and ad hoc meetings were held between students and FAEW partners as needed to achieve individual goals. Total project time was twelve months (July 2020 to June 2021), as stipulated in the President's Challenge award opportunity.

Business Intelligence to Optimize Dashboarding

The volume of data amassed at FAEW over four decades of service to the community contains the essential information needed to support daily, weekly, monthly, and annual decisions that ensure those in need of food are served. Examination of these data in their stored form, or even in a spreadsheet, limited practical value to those leaders needing to make rapid decisions about food acquisition, storage, and distribution activities. Software companies identified this need and created robust platforms for organizations like FAEW to import their data and create visuals, dashboards, and applications. FAEW had a pre-existing license with Tableau; however, this plug-and-play solution has some less obvious embedded functionality that presents data to leaders who wish to utilize their data to gain knowledge. After building a business case with FAEW leadership, students assessed readiness, created a business intelligence roadmap, developed scope, and obtained approvals to optimize current practice.

The main goal of enhancing Tableau practices was to automate reports and save time spent creating dashboards. Historically, FAEW users would export data from Primarius (a food-bank-oriented database management solution) to Excel, replace data sources in Tableau, and perform validation. Students observed visual analytics users at FAEW and found that approximately two hours were spent per day creating these reports. Students then created dashboard templates within Tableau to reduce the time spent repeating these creation steps. The visual component of most of the dashboards did not change; however, the underlying processes were changed to elevate the level of automation as new data were provided. The first two phases allow the newly created workbooks to serve as templates in which only data sources need to be replaced to create new reports. Documentation of changes to the underlying dashboards were documented in one-page guides provided to FAEW containing summary features and narrative text. FAEW representatives with subject matter expertise reviewed and validated all changes. The applications designed by data scientists in training (Marquette students) have been tested and approved by key staff and leaders at FAEW. Subsequent observation of users generating the dashboards revealed that no more than ten minutes were required to generate the same information, corresponding to an estimated 458 hours saved in person-hours per year.

Legacy Data Cleaning with Business Analytics

Preparing clean and well-structured data sets is a prerequisite for reliable and insightful data analytics. FAEW accumulated a large amount of data during its nearly 40 years of operation, providing abundant resources for data analytics. However, raw data usually require cleaning before analytics can be conducted with high reliability and accuracy. Common data aberrancies include incorrect, incomplete, duplicated, or improperly formatted values. The central data components in FAEW's network include (a) agencies (or members), who receive foods from FAEW and distribute them to people in need; (b) donors or vendors, who provide foods or products to FAEW; and (c) products or foods, which flow from donors/vendors to agencies and then to people in hunger, as illustrated in Table 2. The key entities in FAEW's relational database management system (i.e., Primarius) map those major components in their network. Therefore, data cleaning was conducted on data sets for agency, donor/vendor, and product. The fact that FAEW staff were already deeply familiar with data stored within Primarius achieved two goals: FAEW could provide subject matter expertise for data elements and tables to Marquette University students, and there were no barriers to technology adoption.

The data cleaning process included three major steps: data extraction, data analysis, and reporting. Raw data sets were first extracted from Primarius and then imported into RStudio (an Integrated Development Environment for R) for data analytics. Using various R packages (e.g., tidyverse, DT, and knitr), the quality of the raw data sets were examined, and aberrant values (such as missing, incorrect, and redundant values) in the data sets were identified and summarized in reports with visualizations (i.e., charts and tables). Reports were prepared using RMarkdown, which combined texts, codes, and visualizations and improved consistency and reproducibility. The data cleaning process was iterative. After FAEW improved the data quality in Primarius based on the findings in the reports, new data were extracted and sent to the Marquette University team to check the data quality using the same cleaning process, ensuring aberrant data values were fixed. A secure electronic pipeline was explicitly created for this project, and accordant non-disclosure agreements were executed to preserve the confidentiality of FAEW's data. Three data sets (for agency, donor/vendor, and products, respectively) were processed for data cleaning. As anticipated for any large organization with decades worth of data, the agency data set included over 500 aberrant values, involving variables such as agency reference number, agency name, zip code, state, and city. Having clean data subsequently supported the data transformation, visualization, and modeling in other aspects of this project. In addition, the data cleaning process helped FAEW improve the quality of data in Primarius and increase their productivity for various reporting tasks on a daily, weekly, monthly, or annual basis to stakeholders.

Table 2. The number and types of aberrant records, by table, found in FAEW's source data

Table	Number of Aberrant Records	Examples of Variables Impacted
Agencies	560	Agency Ref, Agency Name, City, State, Zip, Contact
Donors and Vendors	962	Donor Ref, Donor Name, City
Products	1,446	Product Ref, Name, Storage

Interacting with Company-Wide Data

Clean data served as the foundation of our team's efforts to extend FAEW's ability to derive meaningful insights and make informed decisions about their food acquisition and distribution practices. An assessment of reports that FAEW leadership created periodically revealed a scarcity of visuals that aggregated data in meaningful ways. Thus, students were assigned to nested groups within FAEW's leadership team to explore the needs of leaders to manage their inventory effectively, interrogate the duration of which products had been stored in two central warehouses, and called-out opportunities to store large, unanticipated food donations rapidly. In addition, students employed human-centered design thinking to support user-driven innovation and technical solutions. Design thinking is taught as part of Marquette University's Visual Analytics course (Department of Computer Science). It follows loosely the framework developed by Stanford University as published in their Design Thinking Bootleg guidebook. A three-page interactive application was subsequently programmed using RShiny, a web application framework for R.

Inventory and Threshold

The inventory and threshold page gave power to the user to investigate quickly and accurately the inventory of specific warehouses through quick-read widgets and a comprehensive data table. The widgets allowed the user to determine the number of on-hand items less than 20 pounds by weight or fewer than five cases, indicating low inventory. Coupled with the widgets was a gauge readout to inform users of a percentage of all inventory items that Wisconsin partner agencies cannot order online. The table allowed the user to search the uploaded data for any insights drawn up outside the widgets and gauge.
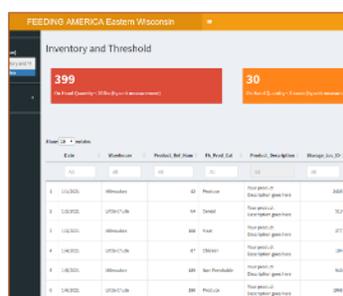
Aging

The aging page gave the investigator quick readouts with widgets of how many items in stock had been in storage for 30, 60, 90, and 120 or more days. This page then listed all relevant items in an easy-to-use searchable table format for more detailed information, such as the warehouse it is in or the product category. This type of information gave agency to the user to make more informed decisions and prioritize tasks and resources to make the most efficient impact on the region's inventory based on time in storage.

Storage Availability

The storage availability page gave the end-user the needed information to instantaneously determine the overall threshold, either too full or not enough, of the organization's warehouses frozen, refrigerated, and dry goods. Three color-coded gauges displayed this information, with < 40% representing green, 40 – 80% representing yellow, and ≥ 80% representing red. This data was also displayed in a time series chart with customizable date and warehouse input. Specifically, the user saw specific moments in time, defined by the user, of the frozen, refrigerated, and dry inventory. This type of information empowered users to see trends in inventory, such as seasonality, that previously might not have been noticed.

Figure 1. Screenshots and brief descriptions of three pages of the interactive web application.


The *Inventory & Threshold* page empowers users to quickly and accurately investigate warehouse inventories, as well as determine the number of on hand items that are less than 20lbs by weight or fewer than five cases. Widget gauge readouts inform the user of the percentage of all inventory items that are not on PWW.
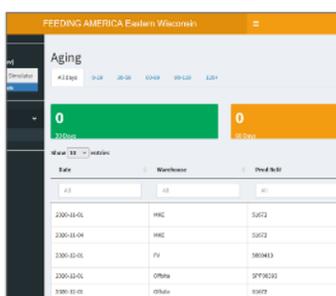
The *Aging* page displays the number of in stock items that have been in storage for 30, 60, 90, and ≥120 days. This page then lists all relevant items in an easy-to-use and searchable table for more detailed information like the warehouse in which the product is located or the product category.
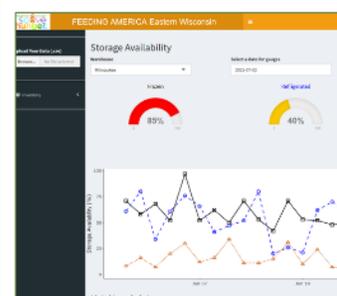
The *Storage Availability* page provides users with needed information to instantaneously determine the overall threshold, either too full or not enough, of the organization's warehouses frozen, refrigerated, and dry goods. This data is displayed in a time series chart with customizable date input.

This tailored, interactive application is available privately to defined users at FAEW through an online web portal. Students created a short vignette of instructions for users to extract, load, and transform a data set from FAEW's source system that consumes fewer than a few minutes. That document was then uploaded into the hosted environment and fed the application with all relevant data. Analysis of improvements in key performance indicators resulting from this application was ongoing.

Attempts to Optimize Purchased Products

As there were gaps between donated food and the food items needed in the community, not all food distributed by FAEW was donated. Therefore, students also attempted to employ predictive analytic techniques, specifically supervised data mining, to aid FAEW in understanding the determinants of purchased products. Possible techniques were evaluated. The Naïve Bayes method for supervised data mining was selected since purchased was a binary outcome variable. Though naïve bayes was efficient and straightforward in terms of computational load, it required categorical predictor variables and made the underlying assumption that all predictor variables were independent. Thus, students interrogated FAEW data (e.g., warehouse, storage, weight, etc.) using classic techniques to confirm variable independence and binned (i.e., grouped) numeric variables as appropriate. More than 20 models were built and interrogated for combined sensitivity and specificity (i.e., if the model is able to predict whether or not a product will be purchased in the future). Unfortunately, even with several environmental variables not contained within the data set itself (e.g., seasonality, region, etc.), model outputs simply lacked the specificity needed by FAEW to make robust predictions about product purchases. Therefore, students are currently partnering with FAEW leadership to conceptualize additional data elements to collect to support more robust analytics.

Transforming a Hunger Needs Formula

Finally, and perhaps one of the more subjective positive outcomes of this collaborative research effort, FAEW leadership was exposed to the skills and talents of data science students. This, combined with FAEW's deep subject matter expertise and wealth of experience in providing food to those in need, created an unanticipated opportunity requested by FAEW leadership. Through a combination of business intelligence, business analytics, and visual analytics, students developed an interactive application that transformed a rudimentary, albeit effective, spreadsheet formula into an interactive, easy-to-use application that allowed FAEW members to input values, such as the number of people in need, the number of meals, the cost per meal, etc., to derive the total number of pounds of food and dollars needed to support a projected increase in need. This fully autonomous application provided FAEW with a sustained tool to ensure that they continue to end hunger in our communities.

Conclusions and Future Directions

Whereas food insecurity decreased in metropolitan areas, food insecurity is increasing in non-metropolitan and rural areas (Nord, 2002). In part due to the SARS-CoV-2 pandemic, it is estimated that food insecurity will affect 1-in-7 Wisconsinites in rural counties during 2021 (Feeding America Eastern Wisconsin, 2021). Unique dietary needs of rural populations and underlying disparities by race, ethnicity, age, and socioeconomic status exacerbate challenges associated with providing quality food to those in need in these rural communities (Jernigan et al., 2017). Foremost among the barriers to feeding food-insecure members of rural communities is a paucity of information about these communities' unique needs and infrastructures. We are in the process of proposing a project that addresses food insecurity in rural places using three phases:

- conduct thorough market research and identify existing data sets to inform strategies to treat food insecurity in rural places, supplementing with data collection as needed,
- utilize statistical modeling, supervised and unsupervised data mining, and forecasting techniques to reveal new insights into the determinants of successful food provision in rural communities, and
- design and implement natural experiments within FAEW's broad food distribution network to test hypotheses that will inform best practices in food distribution across all rural communities.

Lessons learned by students, FAEW leaders, and faculty included an adjustment to entirely remote collaboration due to SARS-CoV-2, the importance of compensating FAEW leaders for buy-out time to divert from daily operations to collaborate with students, an expansion of faculty' office hours to support the development of data-scientists-in-training carefully, and adapting to health and wellness needs during a global pandemic. Setting clear expectations, both program-wide and for each project, helped achieve critical objectives within the short timeline available for funding. In addition, leveraging key personnel's robust project management experience helped address unanticipated challenges with grace. The authors acknowledge that these features may not be routinely available in all academic-community partnerships and encourage a resource assessment before launching such efforts. Much effort focused on bridging the gap between what our community partner needed and what could be achieved by higher education institutions and data scientists in training through various discussions on identifying needs, designing prototypes, using simulated data sets, and testing the applications from users' perspectives. Communication is the key to the success of this collaboration project. Deepening this proven collaborative relationship between Marquette University and Feeding America Eastern Wisconsin will transform work thus far into a larger project with a broader scope to produce meaningful and institutionalized translational research opportunities for years to come.

References

Feeding America Eastern Wisconsin. (2021). *Our impact.* https://feedingamericawi.org/about-us/our-impact/

Jernigan, B. B., Wetherill, M.S., Hearod, J., Jacob, T., Salvatore, A. L., Cannady, T., Grammar, M., Standridge, J., Fox, J., Spiegel, J., Wiley, A., Noonan, C., & Buchwald, D. (2017). Food insecurity and chronic diseases among American Indians in rural Oklahoma: The THRIVE study. *American Journal of Public Health*, *107*(3), 441-446. https://doi.org/10.2105/AJPH.2016.303605

Marquette University (2021). *The president's challenge*. https://www.marquette.edu/innovation/the-presidents-challenge.php

Nord, M. (2002). Rates of food insecurity and hunger unchanged in rural households. *Rural America*, *16*(4), 42-27. https://doi.org/10.22004/ag.econ.289511

U. S. Bureau of Labor Statistics. (2021). *Wisconsin.* https://www.bls.gov/oes/current/oes_wi.htm

U. S. Department of Agriculture. (2021). *Key statistics and graphics.*https://www.ers.usda.gov/topics/food-nutrition-assistance/food-security-in-the-us/key-statistics-graphics.aspx