Predicting Food Bank Demand: A Socioeconomic Analysis and Forecasting Model Investigation

by

Kaitlyn Danielle Lee Rakestraw
B.S. Supply Chain and Information Systems, The Pennsylvania State University, 2019

SUBMITTED TO THE PROGRAM IN SUPPLY CHAIN MANAGEMENT
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF APPLIED SCIENCE IN SUPPLY CHAIN MANAGEMENT
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 2023

Signature of Author: ______________________________
Department of Supply Chain Management
May 12, 2023

Certified by: ______________________________
Inma Borrella, PhD
Research Scientist; Academic Lead, MITx MicroMasters Program
Capstone Advisor

Accepted by: ______________________________
Prof. Yossi Sheffi
Director, Center for Transportation and Logistics
Elisha Gray II Professor of Engineering Systems
Professor, Civil and Environmental Engineering

Predicting Food Bank Demand: A Socioeconomic Analysis and Forecasting Model Investigation

by

Kaitlyn Danielle Lee Rakestraw

Submitted to the Program in Supply Chain Management
on May 12, 2023 in Partial Fulfillment of the
Requirements for the Degree of Master of Applied Science in Supply Chain Management

ABSTRACT

Food insecurity is a problem that affects people in every county within the United States. To combat this issue, many organizations across the country provide charitable food assistance to their communities. The demand for these services is variable, and many of these organizations do not have a consistent method of predicting future demand. This research explores the demand for one food bank, the Mid-Ohio Food Collective (MOFC), and analyzes how this demand differs based on the socioeconomic vulnerability of an area. Additionally, a variety of forecasting models are tested to determine which can best predict demand for the MOFC's services, and the implications of improved forecast accuracy are investigated. This study suggests a framework for categorizing United States counties into two distinct groups, namely "more vulnerable" and "less vulnerable," utilizing socioeconomic factors. When applied to the case study of the MOFC, this classification allowed for identification of differences in demand patterns between the two clusters. Through an evaluation of the RMSE, MAE, and MAPE of nine time series forecasting models, it was found that a naive forecasting model performs well in forecasting demand for the more vulnerable counties in the MOFC's service area. However, it was found that switching from a naive forecasting model to an exponential smoothing model with level and trend components can significantly improve demand forecasting accuracy for the less vulnerable counties. By switching to an exponential smoothing model with level and trend components for the less vulnerable cluster, the MOFC can improve their forecasting accuracy from a 9.9% MAPE using the naïve model to a 4.3%. The factors utilized in this study are relevant and applicable to all counties in the United States. As a result, the insights gained from this research can be effectively employed by food banks throughout the country, enabling them to improve the accuracy of their demand forecasts.

Capstone Advisor: Inma Borrella, PhD
Research Scientist; Academic Lead, MITx MicroMasters Program

ACKNOWLEDGMENTS

I am grateful to Dr. Inma Borrella for her support and feedback throughout the research process. Additionally, I would like to thank Dr. Jarrod Goentzel for his guidance during this process. I would also like to thank the faculty of the MIT Center for Transportation and Logistics for their exceptional teaching and guidance.

Thank you to the Mid-Ohio Food Collective for supporting this project and providing the data that made this research possible. I am appreciative toward each representative from the MOFC who provided valuable insights and answered all the questions that allowed for this research to develop.

I would also like to thank my classmates for their friendship and support. Completing the capstone project alongside this cohort has made it a memorable and meaningful experience.

Finally, I would like to thank my family and friends for their endless love and support.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

In 2021, more than 53 million people in the United States sought food from charitable food organizations (53 Million People Visited Food Banks | Feeding America, n.d.). The need for these services across the United States has been amplified by factors such as rising inflation, unemployment, and the COVID-19 pandemic. 33.8 million people in the United States were living in households considered food insecure in 2021 (USDA ERS - Key Statistics & Graphics, n.d.). The United States Department of Agriculture (USDA) defines food insecurity as a "household-level economic and social condition of limited or uncertain access to adequate food" (USDA ERS - Definitions of Food Security, n.d.).

Many organizations across the United States, from large nonprofit networks to small local food pantries, provide access to the food necessary to maintain a healthy and well-rounded diet to all members of the communities that they serve. A food bank is a large warehouse that receives, stores, processes, and ships food supply out to smaller front-line agencies that serve the community (Crandall, n.d.). These organizations perform and manage a wide breadth of supply chain functions. The nature of food bank operations involves a large amount of uncertainty, including in determining levels of demand over time.

The demand for food banks is variable, changing based on a variety of socioeconomic conditions. Demand visibility can also be a challenge for food banks because they rely on reporting from a large quantity of smaller organizations. This can lead to both overestimating and underestimating their supply needs (Orgut et al., 2016). Food bank operations have traditionally been supply-focused, but recent events, including the COVID-19 pandemic, have demonstrated the need for a more strategic approach to demand forecasting moving forward (Fiocco et al., 2022).

When equipped with more accurate forecasts, food banks can leverage this information to seek additional funding and aim to right-size the amount of food they are securing from donors (Fiocco et al., 2022). Preparing for changes in demand will allow food bank warehouses to properly allocate their supply to frontline organizations, which can increase the number of consumers adequately served, and minimize the amount of supply that is wasted.

The questions to be answered by this research are:

1. How does the socioeconomic status of a county affect demand for food bank services?
2. Which forecasting models can more effectively predict demand for food bank services?
3. What implications would improved demand forecasting have for food bank operations?

The overall goal of this project is to utilize service data from the Mid-Ohio Food Collective (MOFC), a Feeding America member food bank, as well as geographically relevant socioeconomic data, to investigate the demand patterns of this food bank and provide recommendations on how to improve demand forecasting at this food bank. The data will be used to test various demand forecasting models and analyze the quality of forecasts produced by each. Recommendations provided will include the forecasting method to be used and whether this should differ based on the vulnerability of the service area.

The hypothesis is that by segmenting the service area of the MOFC by socioeconomic factors and applying a demand forecasting model, the need of community members served by local food banks can be more accurately predicted. This model could then be used for scenario analyses and could be utilized to advise the MOFC for future decision-making. Another outcome of this project is an analysis of the implications that improved demand forecasting would have for food bank organizations. Improving operations and making more informed decisions surrounding hunger relief could have a significant impact on the success of these organizations.

2. STATE OF THE ART

The areas of literature reviewed to study the problem of improving demand forecasting for food banks are i.) food insecurity and food bank usage in the United States, ii.) food bank supply chain operations, iii.) demand forecasting for food banks along with the factors that affect this demand.

2.1 Food Insecurity and Food Bank Usage in the United States

The USDA defines food insecurity as “a household-level economic and social condition of limited or uncertain access to adequate food.” Additionally, the USDA goes further to define two levels, “low” and “very low” food security. A household having low food security reports having reduced quality, variety, or desirability of diet with little or no reduction in food intake. In contrast, a household with very low food security has multiple indications of disrupted eating patterns and reduced food intake (USDA ERS - Definitions of Food Security, n.d.). The USDA estimates, using data collected by the Bureau of the Census in the Current Population Survey – Food Security Supplement, that 10.2% (13.5 million) of United States households were food insecure at some time during 2021. This percentage of households is a combination of those with low food security (6.4%) and those with very low food security (3.8%) which equates to approximately 33.8 million people living in food-insecure households (USDA ERS - Key Statistics & Graphics, n.d.).

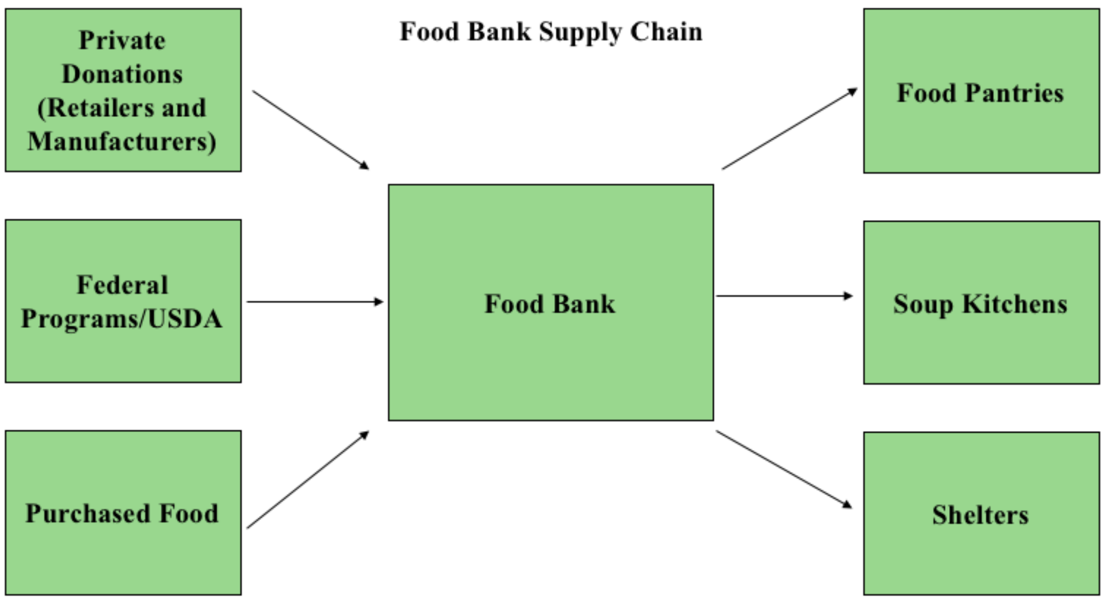
Food banks are non-profit organizations that collect and distribute food to a network of local client-facing charitable food organizations, such as food pantries, that aim to combat food insecurity (How Do Food Banks Work? | Feeding America, n.d.). Feeding America, a nationwide network of around 200 food banks that serve every county in the United States, publishes a yearly report estimating the number of people who have received help from charitable food organizations each year. The Charitable Food Participation Estimate reported that more than 53 million people received help from these organizations in 2021 (53 Million People Visited Food Banks | Feeding America, n.d.). The eligibility requirements for who can receive help from these organizations varies by location but is generally based on household income. For instance, at the Mid-Ohio Food Collective, a Feeding America member food bank, the requirement to receive food is that the household income must be less than 200% of the federal poverty level (*Get Help*, n.d.). At the Greater Pittsburgh Community Food Bank, community members are eligible to receive food if their household income is 150 percent or less of the federal poverty level, or if they are "experiencing an emergency such as a fire or job loss" (Greater Pittsburgh Community Food Bank, 2022).

2.2 Food Bank Supply Chain Overview

Food banks act as warehouses that collect and distribute food to customer-facing hunger-relief organizations within their geographic network. As shown in Figure 1, these warehouses receive, store, process, and ship food to the local partner agencies within their network, which can include food pantries, soup kitchens, shelters, and other charitable food providers (Crandall, n.d.). The food bank receives donations from a variety of sources, and then these goods are inspected and sorted to determine if they are usable. Three primary sources of food for these organizations are private donations, federal programs, and food purchased by the food bank (Morello, 2022). The donations are then allocated into lots to distribute out to the client-facing member organizations in the food bank's service area (Mohan et al., 2010). The specifics of outbound distribution from these food banks vary by location, but it is typically a pull-system. For instance, at the Mid-Ohio Food Collective, the customer-facing agencies place orders through an online inventory system based on what goods they need at their location.

Figure 1

Food Bank Supply Chain



Challenges faced by food bank supply chains include capacity limitations in warehousing and transportation, variability in supply, and uncertain demand (Orgut et al., 2016). Among the 200+ food bank warehouses in the United States, there is considerable variability in physical capacity. Depending on the category of food, the product needs to be stored in either dry, refrigerated, or freezer space. These constraints apply to the agencies that are distributing the food to the community as well – some agencies are very small, and capacity "becomes a primary constraining factor on its ability to store, and hence distribute, food to the food-insecure population in its service region” (Orgut et al., 2016). This capacity challenge was exacerbated by increases in demand and donations as a result of the COVID-19 pandemic onset (Benoliel et al., 2021).

The capacity challenges extend to transportation. Traditionally, the customer-facing agencies (food pantries, soup kitchens, etc.) would come to the food bank warehouse to collect their supply of donations. Food bank warehouses have increasingly been switching to a delivery-based model for distribution (Balcik et al., 2014). This change is due to the fact that many smaller agencies do not own adequate vehicles to transport food in terms of physical size, as well as refrigeration and freezer space.

Another source of uncertainty and challenges for food banks is the variability in both food and monetary supply. The factors that affect availability of supply depend on the source of the donations – for example, donations from retailers are affected by consumer behavior that determines which products have a surplus (Orgut et al., 2016). Grocery chains also donate damaged or close-to-expiration goods, and these quantities can change or be limited by the retailer’s efforts to reduce shrinkage (Mohan et al., 2010). Manufacturers tend to donate products that are safe but have slight defects in appearance that would impact sales in a retail store, and the quantities of these items change based on “adoption of lean production practices along with improvements in manufacturing quality” (Orgut et al., 2016). Donations from local farmers are dependent on the growing season and environment. Additionally, supply from state and federal governments can change based on market conditions and budget decisions (Orgut et al., 2016).

2.3 Food Bank Demand Forecasting

Demand for food banks has been characterized in a variety of ways in past research and by the food banks themselves. These methods of determining demand include calculating the volume of food distributed by the food bank (typically in pounds) and tracking the number of people visiting the client-facing agencies each day. For many food banks, visibility into actual demand is a challenge due to inconsistent or infrequent reporting from the individual front-line agencies. This lack of visibility can lead to food banks overestimating or underestimating their true demand (Orgut et al., 2016).

A study by Tarasuk and Beaton (1999) mentions that food bank demand is generally thought to be correlated directly to the food insecurity levels in a given area. Other previous studies have explored a variety of exogenous factors to determine their effect on these food insecurity levels. A study by Donald Rose published in The Journal of Nutrition published in 1999 utilized data from the Census Population Survey to study the relationship between poverty and food insecurity. This study utilized multivariate logistic regression analysis and found that households in poverty "were >3.5 times as likely to be food insufficient as those with incomes above the poverty thresholds" (Rose, 1999). While this is a strong relationship, the study determined that not every household below the poverty level is considered food insecure, and there are households above the poverty level that show signs of food insecurity. This study also referenced previous work which found that food insecure households were more likely to have experienced unexpected expenses (loss of job, gained a household member, etc.) in the previous year (Olson et al., 1996).

A study completed using data from a Feeding America member food bank in North Carolina aimed to identify economic indicators with a relationship to food insecurity. With quantity of food distributed by the food bank (in pounds) being the dependent variable, the independent variables explored in this study included unemployment levels, enrollment in the food stamp program, the number of meals served by the National School Lunch Program, initial unemployment claims each month, number of people enrolled in Medicaid, and several regionally specific variables. This study utilized stepwise regression analysis to identify significant variables and create a demand forecast (evaluated by R-Squared and adjusted R-Squared values), Principal Components Analysis (PCA), and polynomial models (Okore-Hanson, 2012). An interesting component of this research is that, in addition to comparing these economic indicators for the same time period as the demand distribution period, data on the economic indicators from three, two, and one months prior to the demand distribution period were also studied. This was done to determine whether any variables had a lag effect. The findings of this study were that the different branches within the service area of the food bank responded to variables differently, but that overall, the polynomial forecasts produced the smallest forecast errors (Okore-Hanson, 2012).

A study published in 2016 focused on predicting donor behavior, or the supply-side of food bank operations. This study utilized donation data from a United States food bank and employed a moving average approach, an exponential smoothing method with level and trend (Holt's Method), and an Autoregressive Integrated Moving Average (ARIMA) model (Davis et al., 2016). Another study applied Gaussian Mixture Model (GMM) clustering to a set of food bank demand data, with demand being the number of people visiting the food agencies. They then studied several predictive modeling machine learning techniques using the clustered data. They found that Bayesian Additive Regression Trees (BART) had the best predictive accuracy, which supported the idea that linear models do not capture the complexity of food bank demand data (Sucharitha & Lee, 2022).

Historically, food banks have focused primarily on the supply-side of operations and have distributed any donations that they have been given (Fiocco et al., 2022). In recent years, some food banks have attempted to improve their demand forecasting practices, which vary by organization but in some cases were "limited to yearly reviews using historical, backward-facing data" (Fiocco et al., 2022). The COVID-19 pandemic and resulting surges in demand have been a catalyst for changes in these demand forecasting practices. Second Harvest Heartland, a Feeding America member food bank located in Minnesota, partnered with McKinsey & Company to create a forecasting model that would allow them to plan for future best- and worst-case scenarios in terms of the pounds of food that they would need to distribute to meet demand. This model utilized a formula from Feeding America that connects food insecurity to variables such as poverty and unemployment, in combination with predictions from Oxford Economics on how the COVID-19 pandemic would affect these variables (Fiocco et al., 2022). Having insights into when demand may spike and exceed supply can allow the food banks to take educated action to secure additional funds and donations.

3. DATA AND METHODOLOGY

This research analyzes demand for the MOFC and recommends a forecasting method that can be used by the food bank in the future based on the performance of the models tested. This chapter explains the sources of data used in this analysis, details the methodology used for the clustering analysis to segment the MOFC service area by socioeconomic vulnerability, and describes the forecasting models tested in the study. The data sources section includes descriptions of the datasets used in the analysis as well as how and where these datasets were obtained. The second section focuses on the clustering analysis performed in this research. The purpose of this clustering analysis was to study the hypothesis that the forecasting methodology should differ based on the level of socioeconomic vulnerability in the service area. This section describes the steps taken to prepare the data for analysis, the selection of the appropriate number of clusters, and the algorithm used for clustering. The final methodology section details the forecasting methods that I have tested on the MOFC demand data.

3.1 Data Sources

The primary source of data for this project was the Mid-Ohio Food Collective (MOFC), a Feeding America member food bank. This dataset included the total number of services provided by county, by month, from January 2019 through March 2023. Additional data provided by the MOFC included the total pounds of food distributed by month across their network. Initial analysis of these datasets included identifying trends and potential periods of seasonality within the datasets using descriptive statistics and data visualization.

Public sources of data were used to classify each county in the MOFC's service area by level of socioeconomic vulnerability. As shown in Table 1, information on the median household income and the number of people living under the poverty line were sourced from the United States Census Bureau's Small Area Income and Poverty Estimates (SAIPE) program. Additionally, information on educational attainment for adults aged 25 and older was sourced from the USDA's Economic Research Service. Feeding America's Map the Meal Gap data was used for estimates of the food insecurity rate by county. The Map the Meal Gap datasets include the number of food insecure individuals and the food insecurity rate at both the state and county level in the United States.

Table 1

Variables and Sources

Variable(s)	Source
2019 Poverty Percentage and Median Household Income	U.S. Census Bureau, Small Area Income and Poverty Estimates (SAIPE) Program
2019 Food Insecurity Rates	Feeding America’s Map the Meal Gap - 2021 release, 2019 data
Educational Attainment for Adults Age 25 and Older	USDA, Economic Research Service

Additional insights for this research were derived from interviews with Feeding America member food banks in the United States. The representatives that were spoken to and their respective agencies are listed in Table 2. Questions asked included how operations can differ between these food banks, supply chain challenges that are faced by these organizations, and how they are currently predicting demand for their services. After completion of these interviews, responses were analyzed to identify common patterns and significant differences in the food banks’ operations, specifically within demand forecasting.

Table 2

Interview Log

Food Bank Representative	Agency
CEO/President	Los Angeles Regional Food Bank
SVP of Operations & Senior Project Manager - Data Insights	Mid-Ohio Food Collective
Director of Sourcing and Demand Planning	Second Harvest Heartland (Minneapolis, MN)
CEO	Greater Pittsburgh Community Food Bank

3.2 Clustering

Many socioeconomic factors can be used to characterize geographic regions of the United States including median household income, the number of people living below the poverty line, and the food insecurity rate. While a number of these factors have been explored in past research into food bank demand, such as in Donald Rose's 1999 study that investigated the relationship between poverty and food insecurity, this project took a different approach. An assessment was conducted to cluster counties of the United States based on factors that make these areas more vulnerable and more likely in need of food security assistance. This assessment was conducted utilizing all counties within the United States, and then results were applied to the smaller sample set of 20 counties served by the Mid-Ohio Food Collective. The goal of the clustering exercise was to study the relationship between the vulnerability of a geographic area and the volatility of the area's demand for food bank services.

To understand the differences in the counties served by the MOFC, a k-means clustering approach was used to group together areas that are most similar based on a variety of socioeconomic factors. The k-means clustering approach was chosen due to the ease of experimentation and implementation of this method. The ideal number of clusters was not known prior to this analysis, and k-means clustering allowed for simple analysis to find the desired quantity of clusters. The socioeconomic factors considered in this analysis are the Median Household Income, Poverty Percentage, Food Insecurity Rate, and two education-based factors - the percentage of adults with less than a high school diploma, and the percentage of adults with a bachelor's degree or higher. Data was collected from secondary online sources at a county level for all counties in the United States in order to have a larger sample size for the clustering exercise. 2019 data was used for this analysis due to the fact that this was the most recent year with consistently available socioeconomic data. Additionally, completing this analysis with 2019 data gave a picture of the baseline socioeconomic climate of the MOFC's service area prior to the COVID-19 pandemic affecting these variables.

Before beginning the clustering analysis, a Min Max Scaler was applied to the variables in order to normalize the variables and ensure that each variable is equally scaled and weighted within the clustering analysis. Once the variables were normalized, the k-means clustering analysis was executed several times to find the number of clusters that gave optimal results based on the silhouette score for the clustering analysis. According to this metric, the optimal number of clusters tested was 2.

3.3 Forecasting Models: Time Series Analysis

Based on the literature review, time series analyses have been utilized to predict upstream donor behavior (Davis et al., 2016). This research has expanded upon this by using time series analysis to investigate downstream demand for charitable food assistance. Applying time series analysis provided an indication as to the quality of forecasts that can be created using historical demand data. For this research, Python was used to compare a variety of time series forecasting models, including: simple exponential smoothing, Holt's method with level and trend, exponential smoothing with level and seasonality, the Holt-Winter method with level, trend, and seasonality, Autoregressive Integrated Moving Average (ARIMA) and Seasonal Autoregressive Integrated Moving Average (SARIMA). These models were tested for the full dataset as well as for each cluster to determine if the accuracy of each model differs significantly by cluster.

The quality of all models was evaluated using Mean Absolute Percent Error (MAPE), Mean Absolute Error (MAE), and Root-Mean- Square Error (RMSE). Three error metrics were chosen instead of just one to give a more robust picture of the performance of the models being tested. MAE and RMSE are both scale-dependent metrics, which are useful in comparing models tested on the same set of data. Two of these scale-dependent metrics were chosen due to the fact that RMSE is more sensitive to outliers (Pardalos & Koehler, 2006). MAPE was also included because it is a scale-independent, unit-free measurement of error that allows for comparison for forecast performance across models tested on different datasets.

4. Results

A vulnerability assessment was conducted for every county in the United States which was then applied to the case study for this research, the service area of the Mid-Ohio Food

Collective. The aim of this assessment was to investigate the hypothesis that food bank demand would differ depending on the vulnerability of the area, which, if true, would help the MOFC tailor and improve their forecasting methods for the different areas. To test this hypothesis, demand was analyzed for the MOFC's entire service area, and then by cluster. This allowed for a comparison of demand in areas with different levels of vulnerability, with the goal of identifying any significant differences. Additionally, forecasting models were tested and evaluated for the Mid-Ohio Food Collective service area, and then by cluster, to determine if the models performed differently based on the vulnerability of the area.

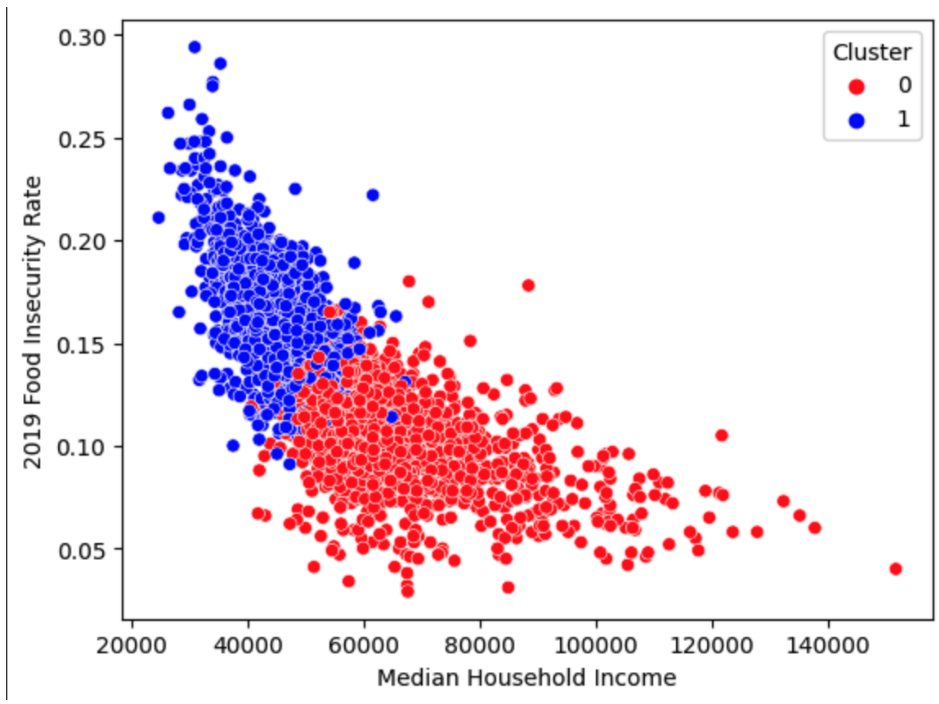
4.1 County-Level Vulnerability Assessment

To evaluate vulnerability across the United States, a k-means clustering analysis was applied to a dataset including socioeconomic metrics for each county in the United States. The metrics used in this analysis were Median Household Income, Poverty Percentage, Food Insecurity Rate, and two education-based factors - the percentage of adults with less than a high school diploma, and the percentage of adults with a bachelor's degree or higher.

The k-means clustering analysis with an n-value of 2 resulted in an average silhouette score of .38. This silhouette score suggests a moderate level of separation and distinctiveness between the clusters. This number of clusters was chosen because the addition of another cluster decreased the silhouette score, indicating a lower quality of clustering analysis. Figure 2 shows an example of these two clusters based on the 2019 Food Insecurity Rate and Median Household Income variables. In this figure, each circle represents a county in the United States, with 3130 in total included in the analysis. Red circles signify counties in Cluster 0, and blue circles are those in Cluster 1. This figure shows each cluster's placement on a graph with the y-axis being the 2019 Food Insecurity Rate, and the x-axis as the Median Household Income.

Figure 2

Depiction of Clusters



As shown in Table 3, the clusters generated by this analysis can be classified as Less Vulnerable (Cluster 0) and More Vulnerable (Cluster 1) based on the socioeconomic variables used in the analysis. Cluster 0 has a lower poverty percent, food insecurity rate, and percentage of adults with less than a high school diploma, and a higher median household income and percentage of adults with a bachelor's degree or higher, which leads to the classification of Less Vulnerable. Cluster 1 is therefore the opposite and classified as More Vulnerable.

Table 3

Cluster Descriptions - All United States Counties

Cluster	Description	Number of U.S. Counties	Average Poverty Percent - All Ages	Average Median Household Income	Average 2019 Food Insecurity Rate	Average % of Adults with Less than a High School Diploma	Average % of Adults with a Bachelor's Degree or Higher
0	Less Vulnerable	1734	10.8%	$64,217	10.6%	8.9%	28.0%
1	More Vulnerable	1396	19.1%	$45,125	16.2%	16.0%	16.8%

The clustering was then applied to this research's case study, the service area of the Mid-Ohio Food Collective (MOFC), a food bank that provides services to 20 counties in the state of Ohio. The 20 counties that are serviced by the MOFC were classified by the cluster that they belong to. This resulted in 9 counties being classified as less vulnerable in Cluster 0, and 11 counties classified as more vulnerable in Cluster 1. Table 4 shows the averages of the metrics used in the clustering analysis for the MOFC service area counties.

Table 4

Cluster Descriptions - MOFC Counties

Cluster	Description	Number of MOFC Counties	Average Poverty Percent - All Ages	Average Median Household Income	Average 2019 Food Insecurity Rate	Average % of Adults with Less than a High School Diploma	Average % of Adults with a Bachelor's Degree or Higher
0	Less Vulnerable	9	9.1%	$72,631	11.3%	8.3%	30.2%
1	More Vulnerable	11	14.4%	$50,388	15.8%	11.8%	15.1%

Table 5 provides a visualization of the clustering application to the MOFC's service area, including the cluster assignment for each county alongside population estimates and the corresponding values for the socioeconomic metrics used in the clustering analysis. This information enables a deeper understanding of the variations in these metrics across the service area, highlighting disparities among different counties and between the clusters.

Table 5

County-level Clustering Details

County, State	Cluster	Population Estimate	Poverty Percent - All Ages	Median Household Income	2019 Food Insecurity Rate	Percent of Adults with Less than a High School Diploma	Percent of Adults with a Bachelor's Degree or Higher
Delaware County, OH	0	220,740	4.8%	$110,252	7.6%	2.8%	57.1%
Fairfield County, OH	0	161,064	8.1%	$71,782	11.0%	6.5%	30.2%
Franklin County, OH	0	1,321,414	13.5%	$64,648	12.8%	8.5%	41.2%
Knox County, OH	0	62,897	11.3%	$59,957	13.2%	8.3%	22.7%
Licking County, OH	0	180,401	9.3%	$66,321	12.4%	7.2%	28.6%
Madison County, OH	0	44,386	9.6%	$65,696	11.6%	11.6%	20.0%
Morrow County, OH	0	35,151	8.5%	$59,498	11.6%	11.0%	15.0%
Pickaway County, OH	0	59,333	11.5%	$63,931	12.9%	12.8%	19.6%
Union County, OH	0	64,971	5.4%	$91,597	8.9%	5.6%	37.8%
Belmont County, OH	1	65,849	11.6%	$50,166	13.9%	8.5%	16.0%
Coshocton County, OH	1	36,618	12.5%	$49,679	14.8%	15.7%	17.1%
Fayette County, OH	1	28,906	13.6%	$51,023	15.7%	15.8%	14.7%
Guernsey County, OH	1	38,287	15.5%	$48,283	17.0%	14.1%	14.8%
Harrison County, OH	1	14,477	14.5%	$50,137	15.4%	11.6%	12.2%
Jefferson County, OH	1	64,789	17.1%	$47,652	16.4%	7.3%	18.3%
Marion County, OH	1	65,291	14.8%	$51,479	15.1%	10.1%	13.1%
Monroe County, OH	1	13,329	14.0%	$48,231	18.1%	10.2%	12.6%
Muskingum County, OH	1	86,408	15.3%	$52,105	15.0%	10.2%	18.2%
Noble County, OH	1	14,176	14.2%	$50,788	15.8%	15.0%	12.1%
Ross County, OH	1	76,891	15.2%	$54,728	16.1%	11.2%	16.9%

Note. Population estimates from Ohio Department of Development (2021). *Ohio Counties 2021 Population and Municipal Population.* https://devresearch.ohio.gov/files/research/P5029.pdf

4.2 Demand Characterization for MOFC Services

Although the clustering analysis was performed for all counties in the United States, the scope of this research is the Mid-Ohio Food Collective's service area in Ohio. The hypothesis is that the vulnerability differences identified in the clustering analysis will impact demand for the MOFC. To begin the analysis of demand, data was analyzed from all 20 counties in order to provide a baseline understanding of the level of food bank demand in the MOFC service area. The graph in Figure 3 shows the demand for services within the 20 counties in the MOFC's service area. The month with the highest number of total services is March 2023, the most recent point in the dataset, and the lowest is January 2019, the first point in the series. Of the 51 periods (months) included in the data, the top 10 months with the highest demand occurred in 2022 or 2023.

Figure 3

MOFC Services Demand

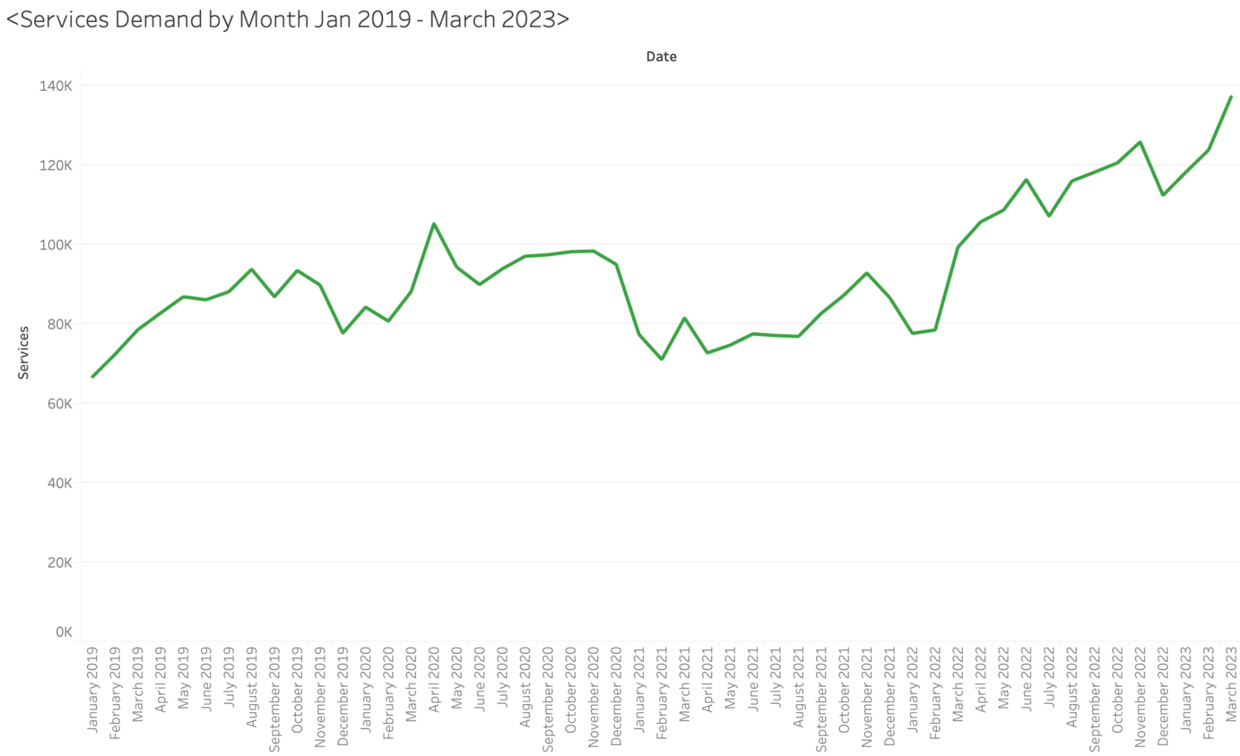
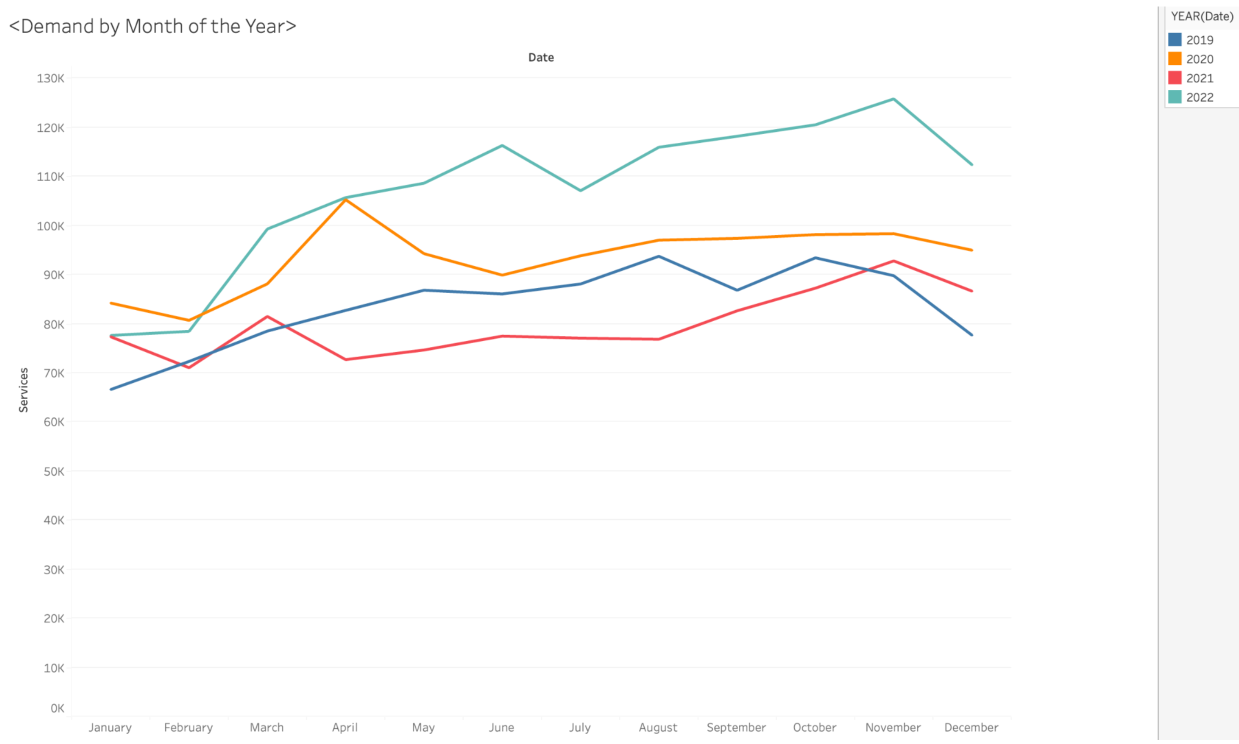


Figure 4 shows the MOFC's demand data by month for each year, excluding 2023 due to incomplete data. In 2021 and 2022, the demand for food bank services peaked in November. Prior to that, in 2020, demand peaked in April - the month following the declaration of COVID-19 as a global pandemic (*WHO Director-General's Opening Remarks at the Media Briefing on COVID-19 - 11 March 2020*, 2020). In 2019, the month with the highest demand was August, closely followed by October. In contrast, the months with the lowest demand for food bank services have consistently occurred in the winter season - in 2019 and 2022, the month with the lowest demand was January, and in 2020 and 2021, it was February. This decrease in demand at the beginning of the year and increase in demand toward the end of the year in November indicates that there may be seasonality in the dataset.

Figure 4

MOFC Demand by Month of the Year



As shown in Table 6, the average number of monthly services for the time period of January 2019 through March 2023 is 93,127, with a standard deviation of 16,393. The coefficient of variation (CV) of this dataset is 0.18, which indicates a low to moderate level of variability within the data. However, from August 2022 through March 2023, demand exceeded one standard deviation from the mean (i.e. demand was greater than 109,520 services) each month.

Table 6

MOFC Demand Data Descriptive Statistics

Average	93,127
Standard Deviation	16,393
Minimum	66,634
Maximum	137,297
Median	89,815
Coefficient of Variation (CV)	0.18

Table 7 breaks down the service area descriptive statistics into a more granular county level. As shown here, there is significant variation in the quantity of service demand by county, with the most populous county of Franklin having an average monthly demand of 60,402, and the county with the lowest demand (Monroe) having an average of only 361. While there is a considerable range in demand between counties, most have similar variability as evidenced by the Coefficient of Variation (CV). The county with the most variation relative to its average and standard deviation is Coshocton County, with a CV of 0.36.

Table 7

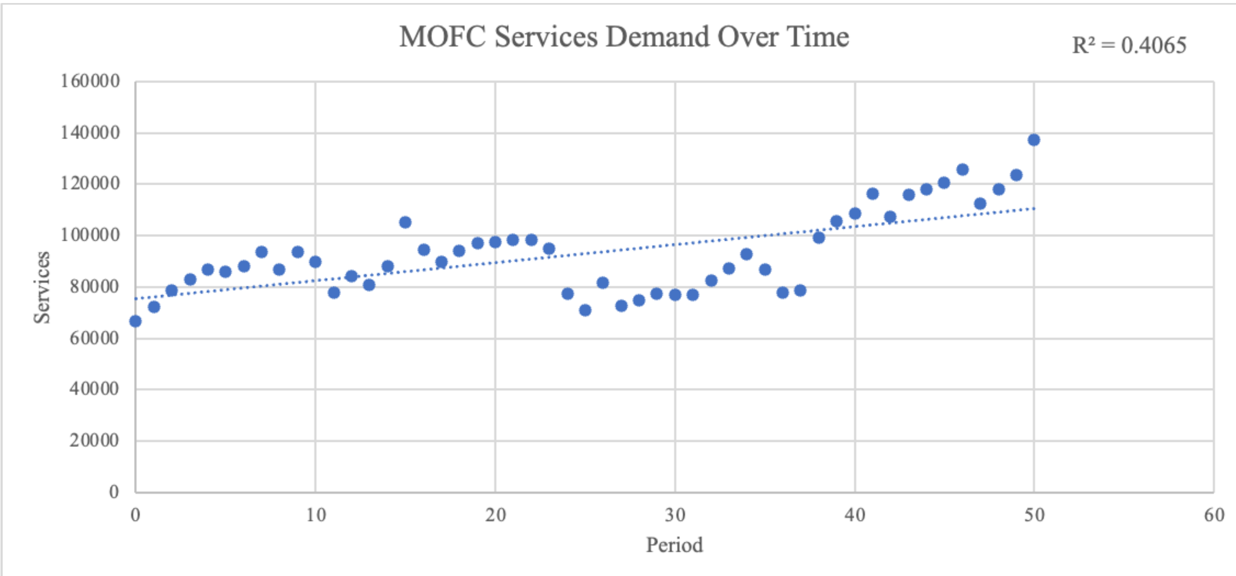
County-level Descriptive Statistics

County, State	Cluster	Average	Standard Deviation	Minimum	Maximum	Median	CV
Delaware County, OH	0	2,218	386	1,464	3,052	2,169	0.17
Fairfield County, OH	0	3,255	686	1,833	4,535	3,235	0.21
Franklin County, OH	0	60,402	13,158	40,269	97,261	56,299	0.22
Knox County, OH	0	1,548	302	1,057	2,172	1,592	0.19
Licking County, OH	0	5,152	806	3,705	7,097	5,184	0.16
Madison County, OH	0	1,091	239	605	1,620	1,142	0.22
Morrow County, OH	0	873	164	558	1,123	878	0.19
Pickaway County, OH	0	1,383	219	1,023	1,907	1,365	0.16
Union County, OH	0	639	210	301	1,071	583	0.33
Belmont County, OH	1	2,419	722	1,229	3,931	2,110	0.30
Coshocton County, OH	1	580	209	260	1,350	543	0.36
Fayette County, OH	1	941	188	550	1,395	947	0.20
Guernsey County, OH	1	1,410	332	775	2,034	1,381	0.24
Harrison County, OH	1	487	97	250	780	467	0.20
Jefferson County, OH	1	2,702	832	1,246	4,374	2,820	0.31
Marion County, OH	1	1,950	452	780	2,935	1,990	0.23
Monroe County, OH	1	361	75	180	523	360	0.21
Muskingum County, OH	1	1,713	347	1,043	2,740	1,686	0.20
Noble County, OH	1	373	99	182	638	370	0.27
Ross County, OH	1	3,628	678	2,467	5,396	3,665	0.19

To analyze the trend of the demand for MOFC services over time, a linear trendline was fit to the data. As shown in Figure 5, the linear trend fits moderately well to the data with an R-squared value of 0.41. This visualization also shows that there may be seasonality in the data, as the data has a pattern of peaks that repeat over time.

Figure 5

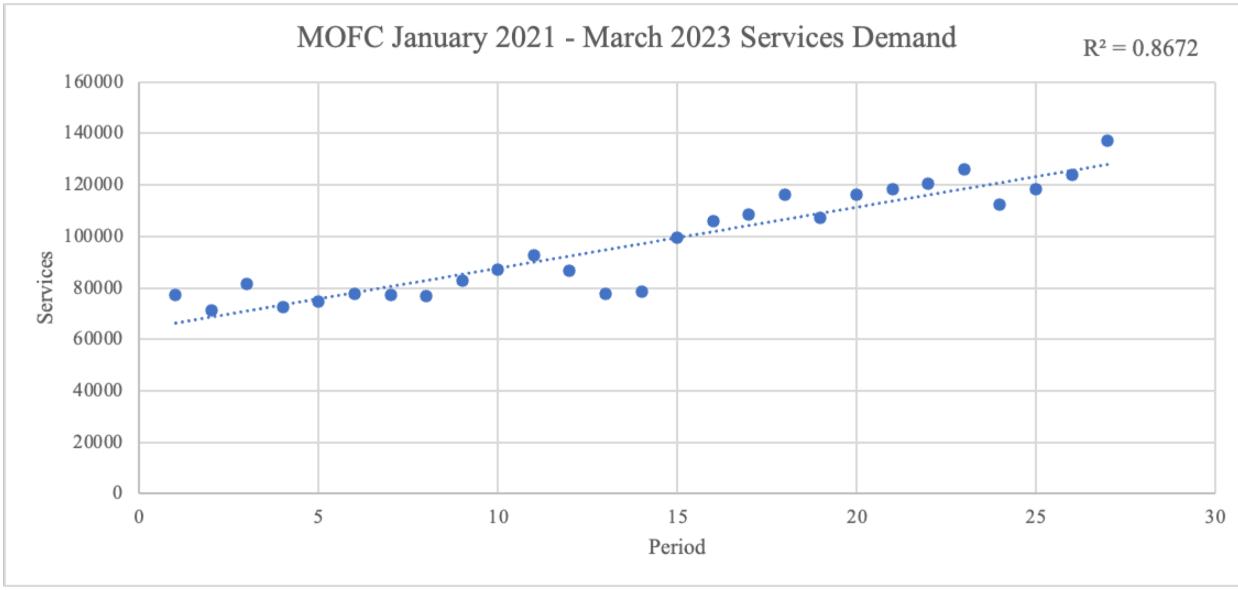
MOFC Services Demand Over Time



Because the initial onset of the COVID-19 pandemic and associated economic effects (for example, a sudden increase in unemployment) occur in the middle of this timeframe, the data from January 2021 through March 2023 was then isolated and the exercise was repeated. This period was chosen as it occurs after the initial effects of the COVID-19 pandemic in 2020. Figure 6 shows that the linear trendline fits significantly better to this isolated time period, with an R-squared value of 0.87.

Figure 6

MOFC January 2021 - March 2023 Services Demand



To continue the analysis corresponding to the hypothesis, the data was then grouped by the cluster that each county in the MOFC service area belonged to in order to analyze each independently. The coefficient of variation (CV) metric was used to compare the demand in the two clusters. It was found that the coefficient of variation for Cluster 0 was 0.19 and for Cluster 1 it was 0.14. Figures 7 and 8 show the demand pattern for each cluster over time.

Figure 7

Cluster 0 Services Demand

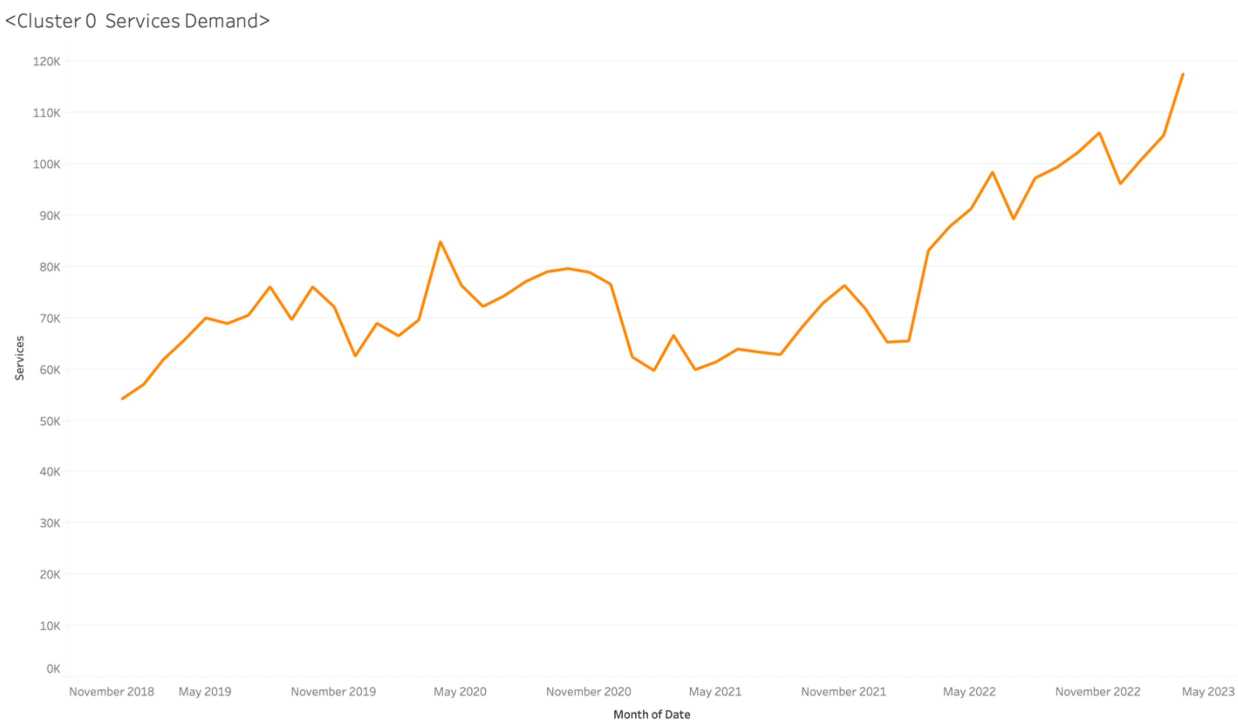


Figure 8

Cluster 1 Services Demand

In order to analyze potential trend and seasonality in the demand data for Cluster 0 and Cluster 1, a linear trendline was applied to the services demand data for each dataset. As shown in Figure 9, the linear trend for Cluster 0 had a similar fit as it did to the full dataset in Figure 5. The R-Squared value of 0.49 shows that this trendline is only capturing 49% of the variation in the data. Because the isolation of data from January 2021 through March 2023 improved the R-Squared value dramatically for the demand dataset of the MOFC's entire service area, the same exercise was repeated for the individual clusters. Figure 10 shows that the trendline for the more recent data of 2021-2023 improves the r-squared value to 0.88.

Figure 9

Cluster 0 Services Demand Over Time

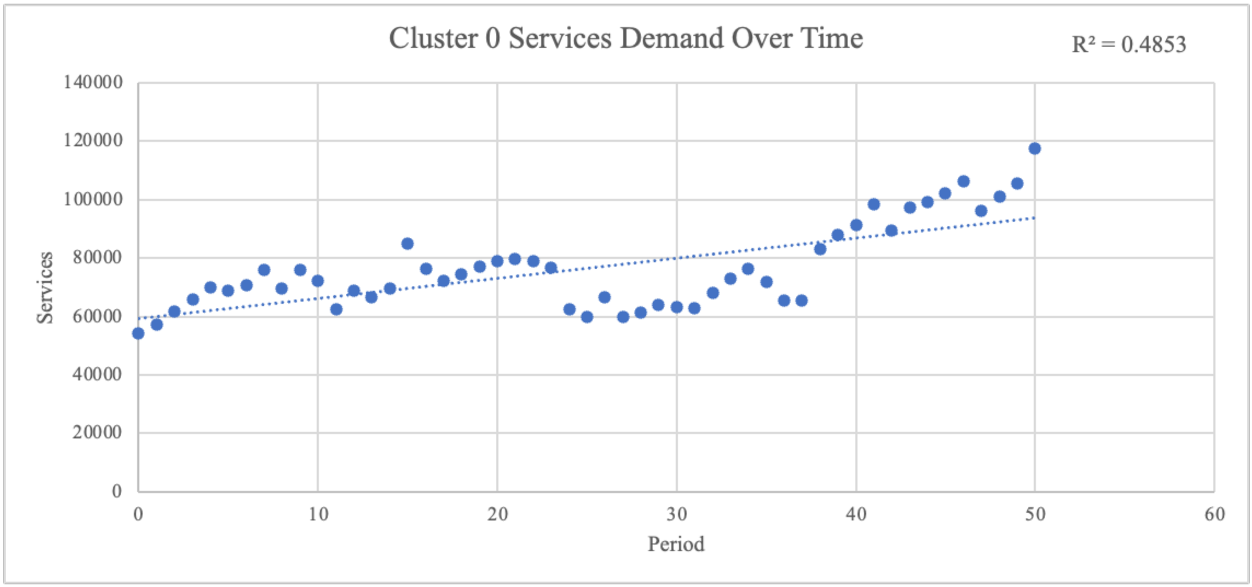
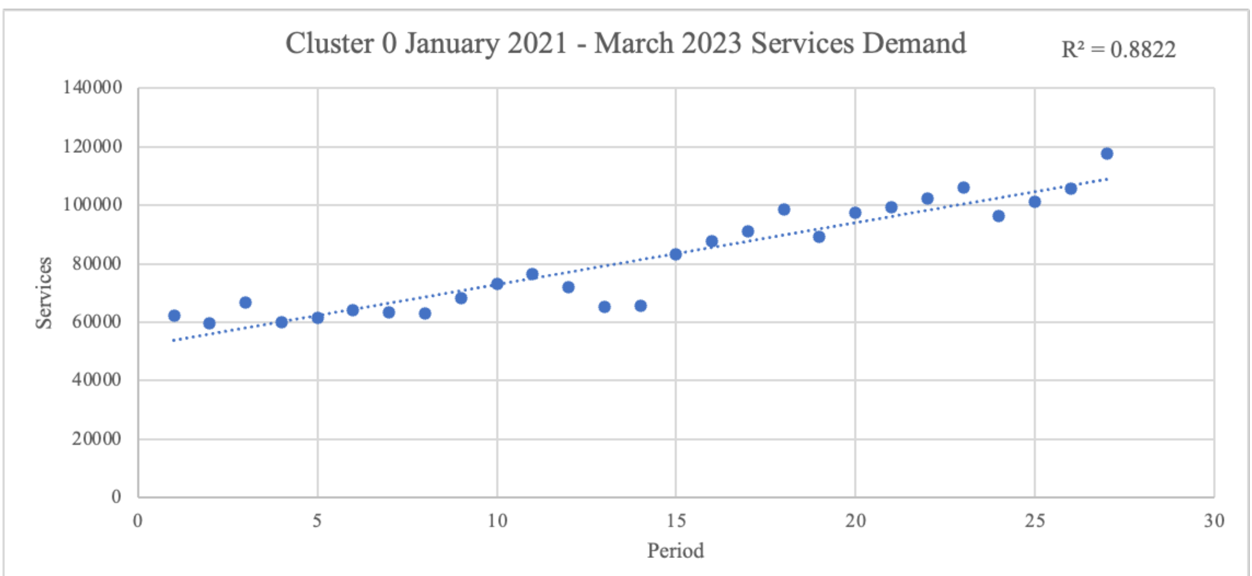


Figure 10

Cluster 0 January 2021 - March 2023 Services Demand



For Cluster 1, shown in Figure 11, the linear trend was a very poor fit to the January 2019 through March 2023 dataset. The R-Squared value of 0.0049 shows that this line captures almost none of the pattern in the data. However, when the trendline was applied to data from January 2021 through March of 2023 for Cluster 1, the R-Squared value improved from almost zero to 0.66, which is shown in Figure 12.

Figure 11

Cluster 1 Services Demand Over Time

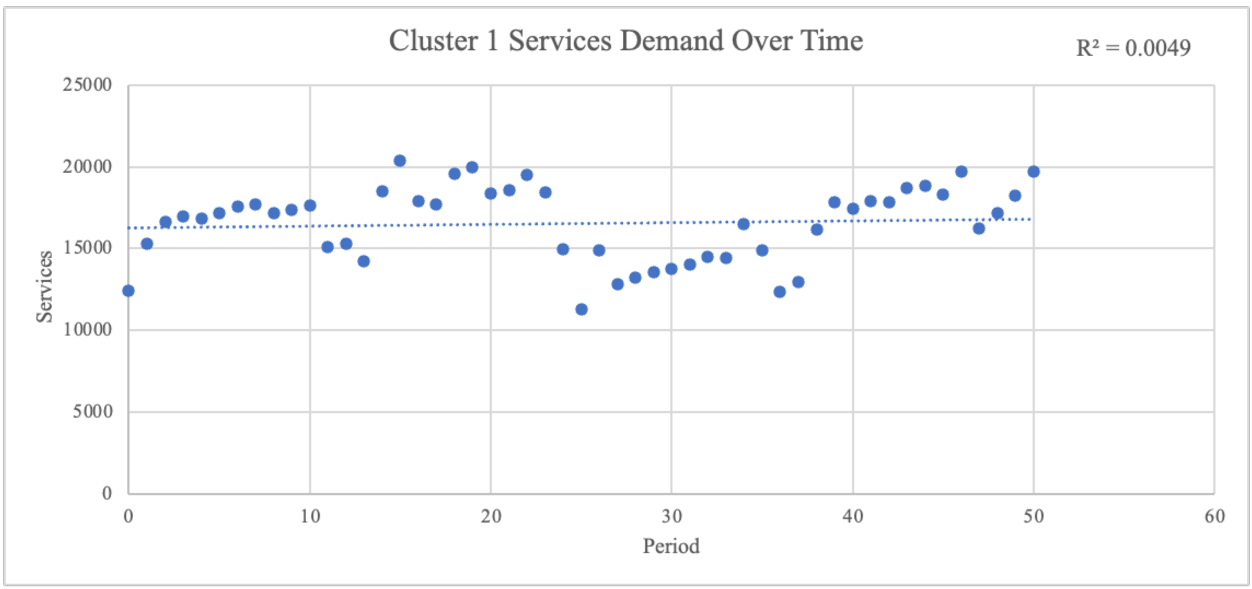
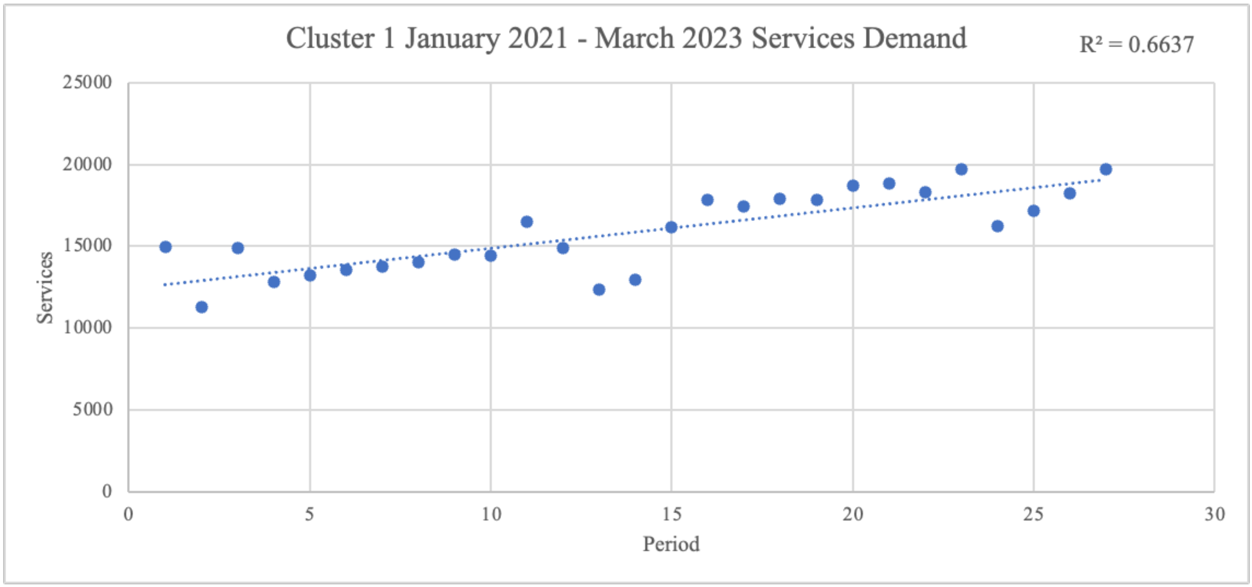


Figure 12

Cluster 1 Services Demand - January 2021 through March 2023



A key takeaway from the demand analysis is the presence of seasonal peaks in demand, particularly during the fall months of October and November. Additionally, the application of a linear trendline to the data suggests that trend has been becoming more impactful in more recent demand observations.

4.3 Time Series Forecasting Models

MOFC Service Area

Utilizing insights from the demand analysis, time series models were chosen to test on the datasets. The first dataset analyzed was the full MOFC service area for the time period of January 2019 through March 2023. For this analysis, 80% of the data was used for training the data (41 periods) and 20% was used for testing (10 periods). As shown in Table 8, the exponential smoothing with level and trend and SARIMA models performed well based on their

MAPE, MAE, and RMSE values. The exponential smoothing with level and trend model was optimized by Python to have an alpha value of 0.995 and beta value of 0.024. This means that the model is placing a low amount of weight on the most recent trend observation and is relying more heavily on the level component. Figure 13 demonstrates the performance of this model in terms of predictions versus the testing dataset. The parameters used by the SARIMA model were values of 1 for autoregressive order (p), differencing order (d), moving average order (q), seasonal autoregressive order (P), seasonal differencing order (D), seasonal moving average (Q), and a seasonal cycle length of 12 to indicate monthly seasonality.

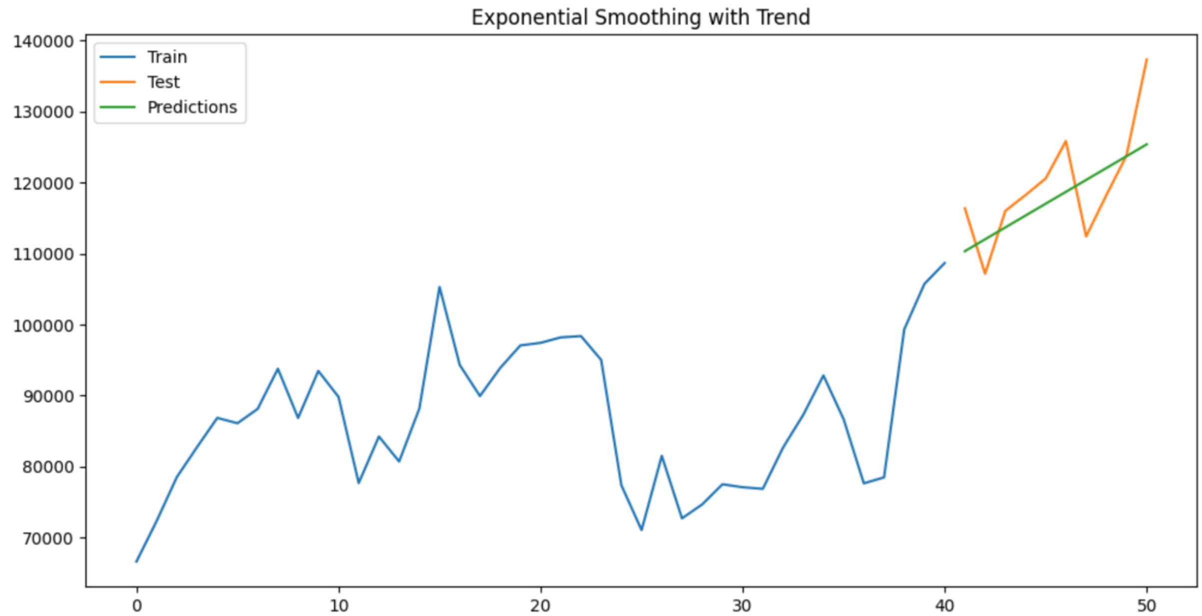
Table 8

MOFC Service Area Forecasting Results

MOFC Service Area			
Model	**RMSE**	**MAE**	**MAPE**
ARIMA	12,974.8	10,749.3	8.7%
SARIMA	8,417.5	6,721.7	5.6%
Simple Exponential Smoothing	13,419.4	11,236.5	9.1%
Naive Forecast	13,407.3	11,224.6	9.0%
Naive Forecast with Seasonality	36,336.4	35,939.8	30.1%
Cumulative Forecast	33,810.2	32,903.2	27.2%
Exponential Smoothing with Level and Trend	5,979.9	5,060.2	4.2%
Exponential Smoothing with Level, Trend, and Seasonality	11,026.2	8,167.3	6.6%
Exponential Smoothing with Level and Seasonality	15,140.1	11,864.5	9.6%

Figure 13

MOFC Service Area - Exponential Smoothing with Level and Trend Predictions vs. Test Data



Cluster 0 - Less Vulnerable Counties

All forecasting methods were also tested on Clusters 0 and 1. As shown in Table 9, the measures of error results are similar for Cluster 0 when compared to the full MOFC Service Area. All parameters remained the same for SARIMA, and the Python model optimized the alpha and beta values for the exponential smoothing model with level and trend to be the same as they were for the full service area (alpha = 0.995, beta = 0.024). Figure 14 shows the predictions of the exponential smoothing with level and trend model with the actualized values of the test dataset.

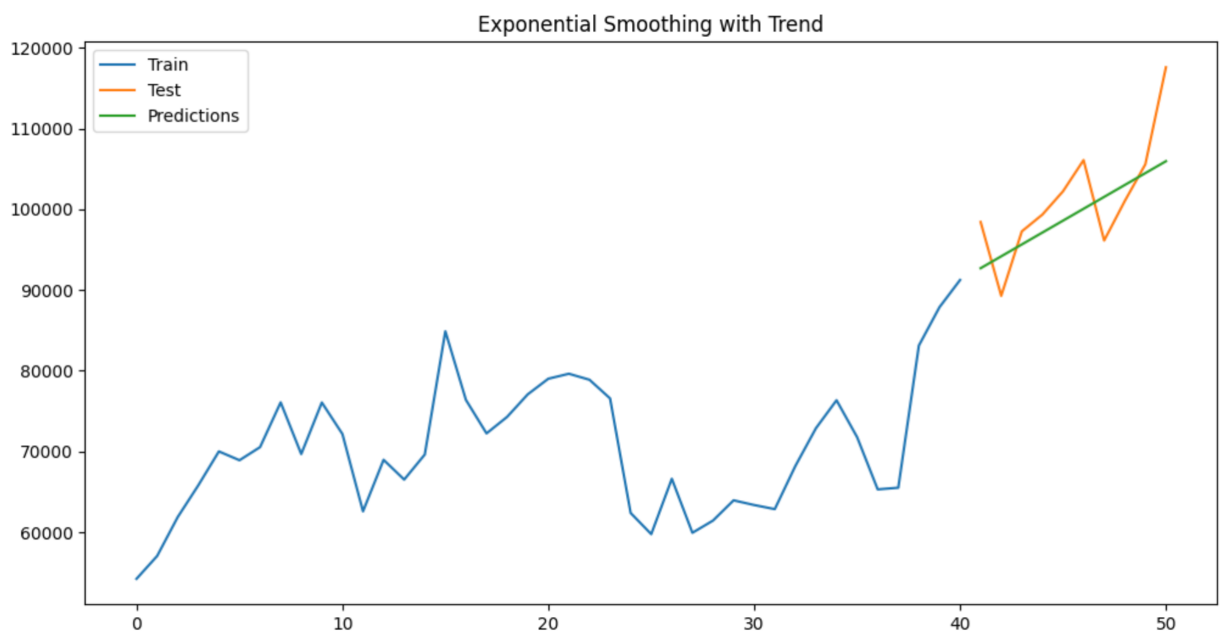
Table 9

Cluster 0 Forecasting Results

Cluster 0			
Model	**RMSE**	**MAE**	**MAPE**
ARIMA	11,845.3	9,879.1	9.4%
SARIMA	5,146.7	4,600.3	4.5%
Simple Exponential Smoothing	12,337.9	10,462.9	9.9%
Naive Forecast	12,324.1	10,449.3	9.9%
Naive Forecast with Seasonality	32,292.4	31,973.3	31.6%
Cumulative Forecast	31,590.0	30,776.5	30.0%
Exponential Smoothing with Level and Trend	5,326.3	4,425.3	4.3%
Exponential Smoothing with Level, Trend, and Seasonality	10,679.3	8,165.2	7.7%
Exponential Smoothing with Level and Seasonality	13,978.7	11,165.0	10.6%

Figure 14

Exponential Smoothing with Level and Trend Predictions vs. Test - Cluster 0



Cluster 1 - More Vulnerable Counties

For Cluster 1, the model that performed the best based on all three error criteria is the SARIMA model. The predicted values from this model are shown in Figure 15. The exponential smoothing with level and trend is still performing well for this cluster, but the parameter weights were optimized to an alpha (level) of 0.995 and trend weight of 0.0001, effectively making this a simple exponential smoothing model. An interesting component of the Cluster 1 analysis to note is that the naive forecast is performing significantly better for Cluster 1 than it is for Cluster 0, with about a 4 percentage point difference in MAPE, which is shown in Table 10.

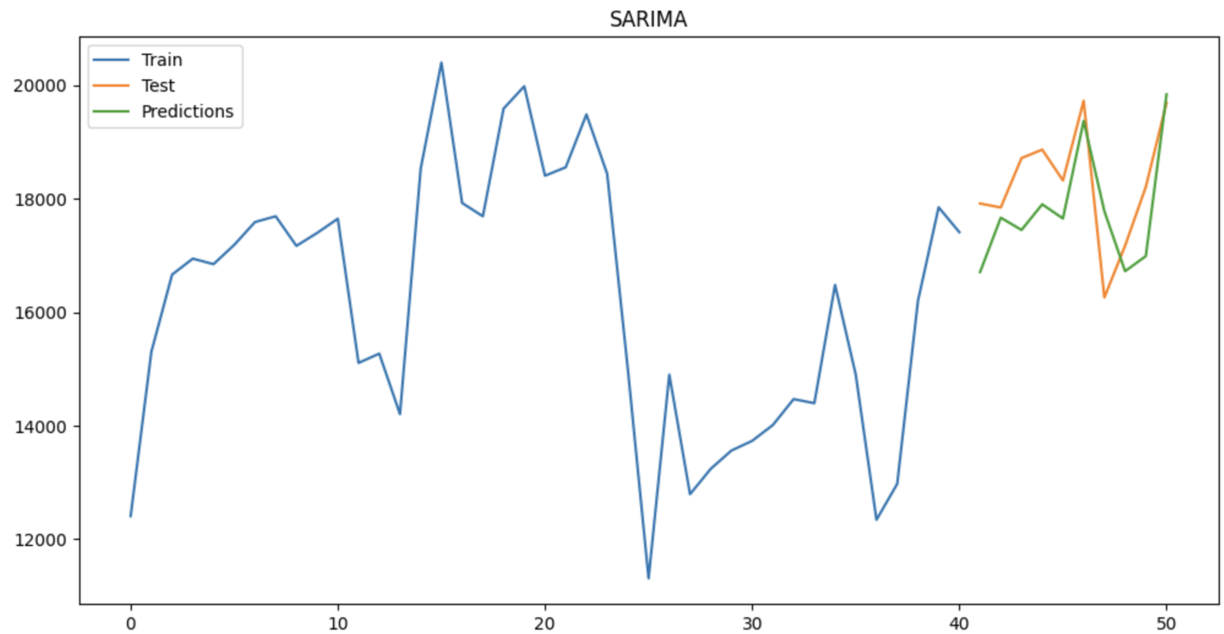
Table 10

Cluster 1 Forecasting Results

Cluster 1			
Model	**RMSE**	**MAE**	**MAPE**
ARIMA	1,389.2	1,194.2	6.4%
SARIMA	928.7	798.3	4.5%
Simple Exponential Smoothing	1,330.1	1,140.1	6.1%
Naive Forecast	1,330.1	1,140.1	6.1%
Naive Forecast with Seasonality	4,100.2	3,966.5	21.7%
Cumulative Forecast	2,355.4	2,126.6	11.4%
Exponential Smoothing with Level and Trend	1,126.4	968.0	5.3%
Exponential Smoothing with Level, Trend, and Seasonality	1,442.5	1,058.2	5.9%
Exponential Smoothing with Level and Seasonality	1,578.4	1,154.6	6.4%

Figure 15

SARIMA Predictions vs. Test Data - Cluster 1



Of all forecasting models tested, the two that showed the most promise with this data are exponential smoothing with level and trend components and SARIMA. For the full MOFC service area, both exponential smoothing with level and trend and SARIMA models performed well in terms of MAPE, MAE, and RMSE values. All forecasting methods were also tested on Clusters 0 and 1. For Cluster 0, the models that performed the best were also SARIMA and exponential smoothing with level and trend. For Cluster 1, the SARIMA model performed the best. The exponential smoothing with level and trend model performed well for Cluster 1, but the parameter weights were optimized to an alpha of 0.995 and a trend weight of 0.0001, making it the same as a simple exponential smoothing model.

5. DISCUSSION

The vulnerability assessment and analysis of demand revealed key insights into the data being used for this research. These insights were the presence of seasonal peaks and declines in the data, and the identification of trend as a potential significant factor in data after 2020. The testing of various time series forecasting models on the data revealed differences in the results between the divisions of the MOFC service area – exponential smoothing with level and trend components performed best for Cluster 0 and SARIMA performed best for Cluster 1. Additionally, it was found that the naïve forecasting model performs better for Cluster 1 than it does for Cluster 0. The results of these analyses support the hypothesis that socioeconomic segmentation and demand forecasting can improve the operations of food banks like the MOFC.

5.1 Vulnerability and Demand Assessment

The clustering analysis provided insights into the vulnerability of different subpopulations within the MOFC's service area of 20 counties. The k-means clustering analysis revealed that a cluster quantity of 2 produced the highest silhouette score of 0.38, indicating a moderate level of distinctiveness between the clusters. Adding another cluster to the analysis led to a decrease in the silhouette score, suggesting that the optimal number of clusters for this analysis is 2. These results suggest that the study area can be broadly classified into two distinct groups based on the socioeconomic factors used in this analysis.

Further analysis of the clusters revealed that Cluster 0 can be classified as Less Vulnerable, while Cluster 1 is More Vulnerable. This classification is based on the socioeconomic factors used in the analysis, which provide insight into the demographic and economic characteristics of each group. The average values for these factors in each cluster also reinforce the distinctiveness of each group - for instance, the average median household income in Cluster 0 is almost $20,000 higher than that of counties in Cluster 1. These results support the fact that there is notable variation in the socioeconomic environment of the MOFC's service area.

5.2 Forecasting Models

The goal of testing a wide variety of time series forecasting models was to identify the most suitable model for each subset of the service area. When analyzing all 20 counties of the MOFC service area, the exponential smoothing with level and trend and SARIMA models performed the best. Interestingly, the optimized beta value for the exponential smoothing with level and trend model was very low at 0.024. Based on discussions with the MOFC and other food banks, it seemed that the model to consider as a current use-case would be the naive forecasting model. The best model, exponential smoothing with level and trend, resulted in a 4.2% MAPE versus the naive forecast which yielded a 9.0% MAPE when applied to the MOFC service area. This equates to an approximate improvement of 53% in MAPE.

Cluster 0 had similar time series forecasting results to the full MOFC service area, with a slight improvement in the performance of the SARIMA model (4.5% Cluster 0 MAPE versus 5.6% for the MOFC service area). The exponential smoothing with level and trend model still outperformed SARIMA, but by a very slim margin. In contrast, SARIMA is the best performing model for Cluster 1, with the lowest RMSE, MAE, and MAPE values of all models tested on that Cluster's demand data. The baseline naive forecasting model performs better for Cluster 1 with only a 6.1% MAPE versus a 9.9% MAPE for Cluster 0. This takeaway suggests that the adoption of a new forecasting technique will have a larger impact on the accuracy of forecasts in Cluster 0.

5.3 Recommendations

A key goal of this research was to identify a forecasting model that can feasibly be repeated by the MOFC and other food banks. Because a naïve forecasting model performed well for Cluster 1 (the more vulnerable counties) with a 6.1% MAPE and is simple to implement, it is recommended to use this model for Cluster 1. For Cluster 0, which contains the less vulnerable counties, the naïve forecasting model did not perform as well, with a MAPE of 9.9%. Due to this, it is recommended for the MOFC to switch to a more sophisticated forecasting model for this cluster. The two forecasting models that performed best for Cluster 0 were exponential smoothing with level and trend and SARIMA. Error metrics between these two forecasting models did not differ dramatically. With this consideration in mind, it is recommended for the MOFC to use an exponential smoothing model with level and trend components due to it being more straightforward and easier to repeat with potentially limited resources. Due to the identification of trend as an influencing factor in recent months, it is recommended for the MOFC to regularly evaluate the performance of the chosen forecasting models to ensure continued accuracy and effectiveness.

5.4 Potential Impact of Improved Forecasting

The implementation of improved demand forecasting at the MOFC could have significant impacts on their operations and ability to plan for future scenarios. Based on comments from members of various departments within the MOFC, having forecasts of the demand for services could have impactful benefits across the organization:

1. Development (Grants): Improved forecasting could help in better understanding the financial resources required to meet the food assistance needs. This would enable the organization to

secure adequate funds for food acquisition and distribution with enough lead time to make changes.

2. Development (Donors/Partners): Forecasting service levels would allow the organization to make a stronger case for support from donors, emphasizing the need for funds to address the projected increase in food distributions. This could also help the organization to secure multi-year commitments from donors to support the anticipated demand.

3. Volunteer Services: Volunteer scheduling can be adjusted up to a year in advance, and improved forecasting could prevent cancellations due to overstaffing or shortages when more volunteers are needed.

4. Agency Services: Enhanced demand forecasting could drive various aspects of food bank operations, such as food sourcing, partner selection, SNAP Outreach staff allocation and deployment, and budgeting decisions, including capacity building fund investments.

5. Communications: More accurate forecasting would influence volunteer recruitment and allocation strategies. Additionally, it would help align public-facing messaging to anticipate future service levels and strategically position the organization's work. Lastly, forecasted service levels can guide communication with elected officials at the local, state, and federal levels, emphasizing the need for and value of the food bank's work.

In summary, improved demand forecasting has the potential to positively impact various aspects of food bank operations. These aspects include financial planning, donor relations, volunteer management, agency services, and communication strategies. These improvements can lead to better resource allocation, increased support from donors, and more effective service delivery for the communities served by food banks.

5.5 Limitations

The limitations of this research include the constrained time frame that was used for forecasting. There seem to have been changes in demand patterns in recent years but incorporating data from years prior to 2019 would provide a stronger basis for training the models and allow for a larger testing set to evaluate the models' performance. A larger dataset would enable the models to better capture long-term trends and historical patterns in food bank demand.

Another limitation of this research is the uniqueness of operations and forecasting capacity for food banks across the United States. Based on insights gathered from the interviews conducted as part of this research, some food banks may not have accurate demand data that captures the true need of their community. The MOFC is at the other end of the spectrum in terms of capabilities, as they have a team of data scientists working to gather insights to help their organization (the largest technical team of any food bank in the United States). The MOFC has a CRM (Customer Relationship Management) system to log the client coming to the pantry or receiving a service from them, allowing for a more robust demand tracking system. This level of data collection might not be present in other food banks, which could affect the ability to apply the suggested forecasting model. The findings from this study are directly applicable to other food banks across the United States if these organizations can collect monthly demand data.

Additionally, the services referenced in this research as demand are not all equivalent – some are a collection of fresh produce, some are a grocery shopping experience, and others are a pickup of food pre-selected by the food bank. This variability in the definition of a service could complicate the model - with the data available for this research, it was not possible to forecast specific types of service. Future research could involve developing separate models for different types of services.

5.6 Future Research

Future research in studying food bank demand could involve a more granular analysis of the demand patterns by exploring weekly and daily data. This approach will enable researchers to identify trend and seasonality patterns that may not be apparent when looking at monthly data. By understanding these shorter-term fluctuations, food banks would be positioned to more strategically allocate their resources and improve their overall operations.

Future research could benefit from exploring exogenous factors in a multiple regression model to determine whether variation in demand can be captured by the model. Some potential factors for analysis include economic indicators such as unemployment and inflation, demographic variables, holidays, usage of other food assistance programs, and weather. Using a multiple regression model to incorporate these factors could allow for further understanding of which factors are correlated with fluctuations in food bank demand.

6. CONCLUSION

Food insecurity is a significant issue in the United States, with many individuals and families seeking support from charitable food organizations. To combat this issue, food banks across the United States, including the Mid-Ohio Food Collective (MOFC) play a crucial role in providing necessary food assistance to affected communities. However, these organizations face the challenge of demand variability, emphasizing the need for accurate forecasting to optimize their operations.

Through this research, the level of socioeconomic vulnerability throughout the MOFC's service area was studied using a k-means clustering analysis. This segmentation was used to examine the performance of various time series forecasting models on each cluster, as well as on the full MOFC service area. Additionally, the implications of improved demand forecasting on food bank operations were examined, specifically from the perspective of the MOFC.

The research findings suggest that by applying appropriate demand forecasting models based on the socioeconomic vulnerability of the area, the needs of community members served by local food banks can be more accurately predicted. For Cluster 1, including the more vulnerable counties, a naïve forecasting model was substantially effective, achieving a MAPE of 6.1%. However, for Cluster 0, which includes the less vulnerable counties, and aggregated forecasts for both Clusters, it is recommended to switch to the exponential smoothing with level and trend model due to the significant improvement in MAPE versus the naïve model. Switching from the naive model to an exponential smoothing model with level and trend components improves the MAPE from 9.9% to 4.3% for Cluster 0.

Improved forecasts would allow the MOFC to inform decision-making and resource allocation to ensure that they are meeting the food assistance needs of their communities. More accurate forecasts would enable the organization to secure the right amount of food supply and distribute it strategically to frontline agencies, minimizing waste and maximizing service to those in need.

In summary, this research supports the hypothesis that socioeconomic segmentation and demand forecasting can improve the operations of food banks like the MOFC. While there are challenges posed by demand variability and unpredictable socioeconomic factors, the application of appropriate forecasting models and strategic planning can significantly improve how food banks address food insecurity.

REFERENCES

53 Million People Visited Food Banks | Feeding America. (n.d.). Retrieved October 13, 2022, from https://www.feedingamerica.org/about-us/press-room/53-million-received-help-2021

Benoliel, E. T., Calvo, A. P., Felix, I., & Ruby, B. (2021, February 19). *Transforming food banks' warehouse operations to improve critical performance*. McKinsey & Company. https://www.mckinsey.com/featured-insights/food-security/transforming-food-banks-warehouse-operations-to-improve-critical-performance

Crandall, D. (n.d.). *A Food Bank Supply Chain: Logistics of Feeding the Hungry | Bastian Solutions*. Retrieved October 17, 2022, from https://www.bastiansolutions.com/blog/a-food-bank-supply-chain-the-logistics-of-feeding-the-hungry/

Davis, L. B., Jiang, S. X., Morgan, S. D., Nuamah, I. A., & Terry, J. R. (2016). Analysis and prediction of food donation behavior for a domestic hunger relief organization. *International Journal of Production Economics*, *182*, 26–37. https://doi.org/10.1016/j.ijpe.2016.07.020

Fiocco, D., Lucas, A., Knoch-Dohlin, M., & Ruby, B. (2022, August 1). *Making a difference with data: What it can mean for food banks*. McKinsey & Company. Retrieved October 20, 2022, from https://www.mckinsey.com/featured-insights/food-security/making-a-difference-with-data

Get Help. (n.d.). https://mofc.org/get-help/

Greater Pittsburgh Community Food Bank. (2022, November 9). *Get Help With Food*. https://pittsburghfoodbank.org/find-food/get-help/

How Do Food Banks Work? | Feeding America. (n.d.). https://www.feedingamerica.org/our-work/food-bank-network

Mohan, S., Gopalakrishnan, M., & Mizzi, P. J. (2013). Improving the efficiency of a non-profit supply chain for the food insecure. *International Journal of Production Economics*, *143*(2). https://doi.org/10.1016/j.ijpe.2011.05.019

Morello, P. (2022, April 20). *How food banks and food pantries get their food*. Feeding America. https://www.feedingamerica.org/hunger-blog/how-food-banks-and-food-pantries-get-their-food

Ohio Counties 2021 Population and Municipal Population Ratio. (n.d.). Ohio Department of Development. Retrieved May 9, 2023, from https://devresearch.ohio.gov/files/research/P5029.pdf

Okore-Hanson, A. (2012). *Empirical Modeling of Food Demand and Characterizing Donor Behavior During a Natural Disaster* [Master's Thesis]. North Carolina Agricultural and Technical State University.

Olson, C. M., Rauschenbach, B. S., Frongillo Jr., E. A., & Kendall, A. (1996). *Factors Contributing to Household Food Insecurity in a Rural Upstate New York County*. Institute for Research on Poverty, University of Wisconsin-Madison.

Orgut, S. I., Iii, L. B. G., Davis, L. B., Ivy, J. S., Jiang, S., Morgan, S. D., Uzsoy, R., Hale, C., & Middleton, E. (2016). Achieving Equity, Effectiveness, and Efficiency in Food Bank Operations: Strategies for Feeding America with Implications for Global Hunger Relief.

In C. Zobel, N. Altay, & M. Haselkorn (Eds.), *Advances in Managing Humanitarian Operations. International Series in Operations Research & Management Science.* Springer, Cham. https://link.springer.com/chapter/10.1007/978-3-319-24418-1_11?

Pardalos, P. M., & Koehler, A. E. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting*, *22*(4), 679–688. https://doi.org/10.1016/j.ijforecast.2006.03.001

Rose, D. (1999). Economic Determinants and Dietary Consequences of Food Insecurity in the United States. *The Journal of Nutrition*, *129*(2), 517S-520S. https://doi.org/10.1093/jn/129.2.517s

State and U.S. Population Estimates. (n.d.). https://devresearch.ohio.gov/reports_pop_est.htm

Sucharitha, R. S., & Lee, S. (2022). GMM Clustering for In-depth Food Accessibility Pattern Exploration and Prediction Model of Food Demand Behavior. *Cornell University - ArXiv*. https://doi.org/10.48550/arxiv.2202.01347

Tarasuk, V. S., & Beaton, G. H. (1999). Household Food Insecurity and Hunger Among Families Using Food Banks. *Canadian Journal of Public Health*, *90*(2), 109–113. https://doi.org/10.1007/bf03404112

USDA ERS - Definitions of Food Security. (n.d.). Retrieved October 17, 2022, from https://www.ers.usda.gov/topics/food-nutrition-assistance/food-security-in-the-u-s/definitions-of-food-security/

USDA ERS - Key Statistics & Graphics. (n.d.). Retrieved October 13, 2022, from https://www.ers.usda.gov/topics/food-nutrition-assistance/food-security-in-the-u-s/key-statistics-graphics/

WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020. (2020, March 11). https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020