Proceedings of the 2019 IISE Annual Conference
H.E. Romeijn, A Schaefer, R. Thomas, eds.

Effect of Underlying Factors on Food Distribution Forecasts Using Visual Analytics

Abstract ID: 578670

Kehinde Odubela, Steven Jiang, Ritson Delpish and Lauren Davis
Department of Industrial & Systems Engineering
North Carolina A & T State University
Greensboro, NC 27411, USA

Abstract

Non-profit hunger relief organizations rely on the goodwill of donors for their in-kind cash, food donations and other supplies to alleviate hunger, reduce human suffering and save lives. However, these organizations struggle with changing demand and supply patterns, disruptions caused by very low donations even though they must make strategic distribution decisions. Food distribution forecasts based on times series models can be useful for these decisions. Yet, it is plausible that food distribution by hunger relief organizations (and demand by the people in need) are driven by certain underlying factors. In this research, we used Visual Analytics (VA) to study the effect of certain underlying factors on the forecast generated for food distribution to the aid recipients. Specifically, we used already tested forecasting techniques to predict the expected quantity of distributed food for the underlying factors identified.

Keywords
Visual analytics, Distribution forecast, Hunger relief

1. Introduction

Individuals and households experience food insecurity when they have limited access to safe and nutritious food they need for an active, healthy life [1, 2]. The Economic Research Service (ERS) of the United States Department of Agriculture (USDA) estimated only 88.2 percent of U.S. households were food secure throughout the entire year in 2017 [3]. Thus, 11.8 percent of households were food insecure at least some time during the year. These include 4.5 percent that experienced very low food security, that is, one or more times, the food intake of the household members was reduced, and their eating patterns were disrupted. This is mostly because they lacked money and other resources for obtaining food [3].

Humanitarian relief organizations such as food banks play a significant role in the fight against hunger. Food banks are nonprofit organizations that serve people who are unable to consistently access enough nutritious food to live a healthy, active lifestyle. Food banks receive, sort and distribute millions of pounds of nutritious food to people in need through a network of partner agencies and direct distributions.

In this research, we worked with a local food bank, the Food Bank of Central and Eastern North Carolina (FBCENC). FBCENC is one of the seven food banks in North Carolina, all belonging to the Feeding America Network – America's largest hunger-relief organization. FBCENC serves about a third of all counties in North Carolina, 34 in total (more than 3.7 million individuals). The counties are reached through 6 food bank branch locations (warehouses) based on the service area under which they are categorized. FBCENC and its branch warehouses can receive donations from wholesale grocers, supermarkets, manufacturers, farmers and other individuals or groups. These donations are distributed through more than 800 partner agencies offering different programs such as kids' café, soup kitchens, children and elderly nutrition, food pantries and so on. FBCENC distributed 70 Million pounds of food in 2016-2017 alone [4]. The operating environment for the food banks such as FBCENC has become increasingly complex. They have dynamic distribution networks with multiple configurations necessitated by the many charitable agency partners that can receive donated food. Food Banks face a significant level of uncertainty relative to both supply and demand which increases the difficulty in understanding available supply and food demands [5]. The reliance on donation further complicates the classic supply chain dilemma to satisfy demand with adequate supply. Supply uncertainty is a product of unknown frequency, amount and quality of items donated well in advance. Demand uncertainty also occurs because expected food need is dependent on complex factors related to poverty and unemployment [6].

Food banks often rely on forecasts to assist their decision making. Forecasting for supply or demand can be done using time series models. Time-series analysis and forecasting take advantage of the past behavior of a sequence of events or measurements to derive future predictions [7]. This analysis also assumes the observed values follow a specific pattern over the period considered, even though it might include random noise which increases the difficulty of identifying the unnoticeable pattern. These patterns are mostly composed of trend and seasonality and can be extended to future periods to generate forecasts of new measurements [8]. Commonly used models for time-series forecasting include Exponential Smoothing (EM), Autoregressive Integrated Moving Average (ARIMA) and their variants (e.g., the Holts-Winters Method).

Previous research on forecasting in-kind donations for the food banks using time series models have been proven to be effective [6]. However, the forecast can be improved by identifying other underlying measures (e.g., socio-economic factors) that influence or have a relationship with the measure of interest. The forecasting model becomes stronger and better captures reality when these underlying measures are introduced. Moreover, it remains a challenge to present the analytical results to food bank operations managers and more importantly allow them to interact with the analytical results. To address these concerns, Visual Analytics (VA) was used in this research. VA bridges this gap between the business and analytics domains and provides a fast and convenient approach for business leaders and decision makers to generate insights from the data without requiring knowledge of forecasting techniques or other algorithms. As a result, food bank decision makers can apply their knowledge of the operations of a hunger-relief organization coupled with the interactive and visual interfaces generated by the VA tool. They can quickly identify patterns and trends in their data and what underlying factors have a relationship with certain parameters.

2. Method

2.1 Data Collection

We obtained five fiscal years of distribution data (from 2006 to 2011) from FBCENC. A fiscal year starts from July and ends in June of the following year. The data is pre-processed to include only food distribution to each of the 34 counties. *Gross weight* refers to the quantity (in pounds) of food distributed per transaction and represents the dependent variable to be forecasted. One of our goals is to determine which model generates the most accurate forecast per month, therefore, the historical data is aggregated by month and year from the *posting date* field. We then obtained historical records of several socio-economic factors that we suspect could have a relationship with the data. We obtained local unemployment statistics from the United States Bureau of Labor Statistics (BLS) at the county level (not seasonally adjusted) and state level (seasonally adjusted) from 2006 to 2011 to match the months and years of data we obtained from the food bank. The unemployment rates are available for each month for every year considered. We also obtained data for poverty rates (for different age groups – Under Age 18, Ages 5 to 17 and Under Age 5) and median household income for 2006 to 2011 from the Small Area Income Poverty Estimates (SAIPE) page of the United States Census Bureau (USCB) website. A description of each measure is summarized in Table 1.

2.2 Building Time Series Models

We chose SAS Visual Analytics 7.4 (SAS VA) in this research because of its ability to deal with huge data sets, automatically select the best chart (auto-charting) for visualizing the data and its advanced forecasting features. The forecasting models currently available in SAS VA include ARIMA, damped trend exponential smoothing, seasonal exponential smoothing, simple exponential smoothing, linear exponential smoothing and Winters method. SAS VA allows the user to manipulate the forecasted data to gain insights into how certain data items factor into the forecast. This involves finding underlying factors that have a relationship to the forecast which can be manipulated with scenario analysis. SAS VA forecasting models can include other measures when it performs the analysis to generate the forecasts. The software maintains the forecasting model that has the best fit with the data.

Table 1: Data Summary

Measure (State and County Level)	Source	Measurement Unit	Aggregation	Period Considered
Food Distribution - Gross Weight	FBCENC	Pounds		July 2006 - June 2011
Unemployment Rate	BLS	Percentage	Monthly	July 2006 - June 2011
Poverty Rates	USCB			
Under Age 18		Percentage	Yearly	2006 - 2011
Ages 5 - 17		Percentage	Yearly	2006 - 2011
Under Age 5		Percentage	Yearly	2006 - 2011
Median Household Income	USCB	Dollars	Yearly	2006 - 2011

3. Results and Analysis

We first investigated the relationship among food distribution and the underlying factors. We then used appropriate times series model for the distribution data at both the foodbank level and county level as well as with and without underlying factors using SAS VA.

3.1 Correlation of Food Distribution and Underlying Factors

We conducted correlation analysis to understand the relationship among the gross weight of food distributed ('Distribution') and the underlying factors studied – poverty, unemployment and median household income. As seen in Figure 1, all variables have a strong relationship (r > 0.8) with food distribution except median household income that has a moderate relationship (r=0.4). Median household income also had moderate relationship with other underlying factors considered.

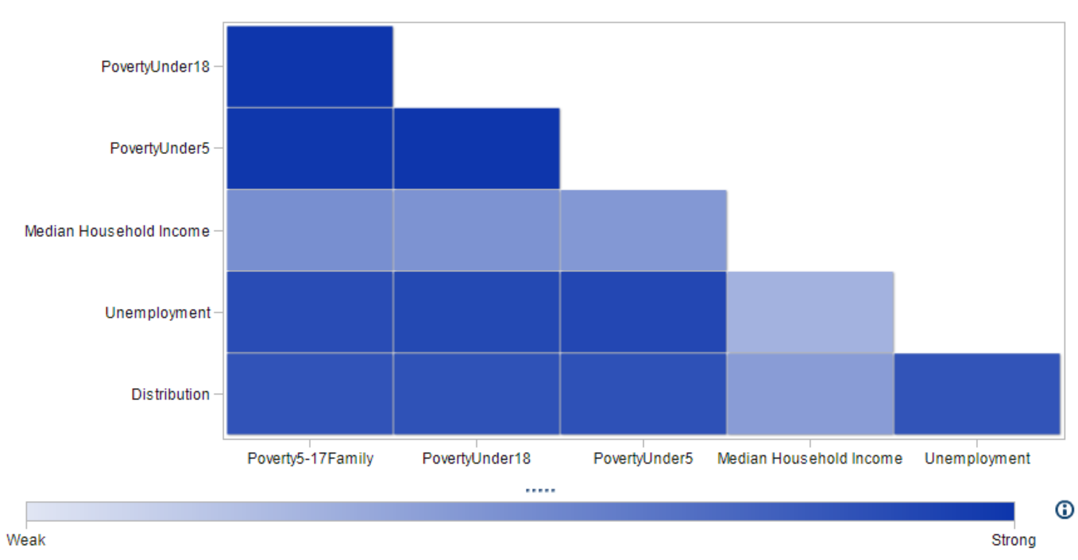

Figure 1: Correlation analysis among food distribution and underlying factors

3.2 Forecast Model Results at the Foodbank Level

Before the introduction of the underlying factors, SAS VA identified Additive Winters Method as the best fit for the data as seen in Figure 2. This is consistent with the findings of Chen (1997) that highlighted the robustness of the Holt-Winters method - its stable forecasting performance and satisfactory forecasting accuracy [9]. Holts-Winters method is especially useful for wide ranging time series that have stochastic or deterministic trend and seasonal components that are subject to structural changes. After the underling factors were introduced to the time series, the forecasting algorithm was updated to the ARIMA model for generating the most accurate forecast for the time series as seen in Figure 3. Among the many underlying factors introduced, SAS VA identifies those that have any contribution to the forecast. At the foodbank level (34 counties), under-18 poverty was identified as the only contributor to the forecast despite the presence of other factors like unemployment and median household income.

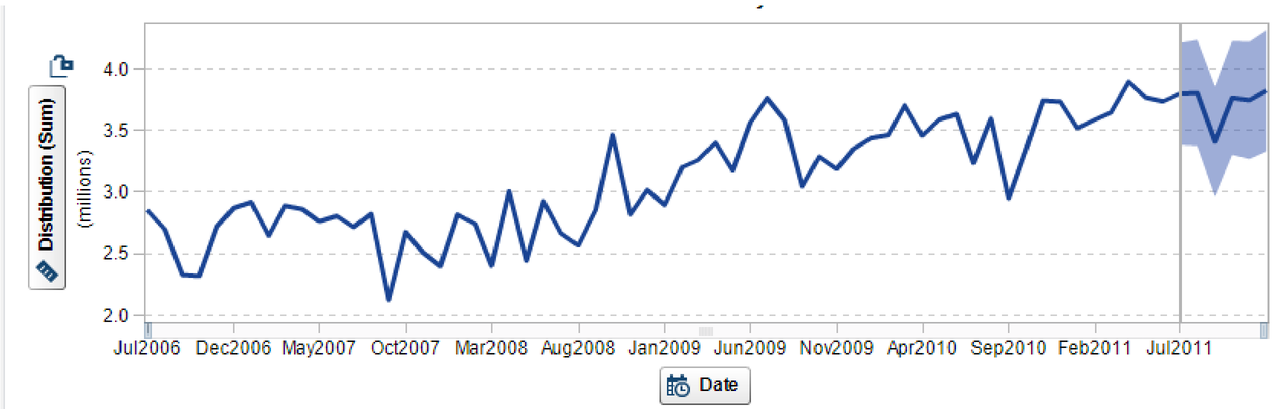

Figure 2: Forecast of monthly food distribution – before introducing underlying factors

3.3 Forecast Model Results at the County Level

The analysis above was repeated using data for one of the thirty-four counties served by food bank. We chose a county instead of one of the six branches served by the foodbank since it was easier to obtain data for the socio-economic factors we studied from publicly available sources at the county level. Wake County is the largest of the 34 counties served by the food bank and the second most populous county in the state [10].

Figure 4 shows the time series plot and forecast of the monthly gross weight of food received by Wake County. SAS VA identified the Damped Trend Exponential Smoothing model as the best fit for the data. Unlike the forecasts generated by the Holts-Winters Method which indefinitely projects the trend in the data, the Damped Trend Exponential Smoothing method dampens the trend to a flat line for the forecast. Gardner & McKenzie (1985) introduced the dampening parameter to avoid over-forecasting the trend particularly for longer time horizons [11]. Forecast methods that dampen the trend have been identified to be very successful when predictions are generated automatically for several series [12]. After the underling factors were introduced to the time series, the forecasting algorithm was updated to the ARIMA model for generating the most accurate forecast as seen in Figure 5. SAS VA identified unemployment and under-18 poverty as significant contributors to the forecast.

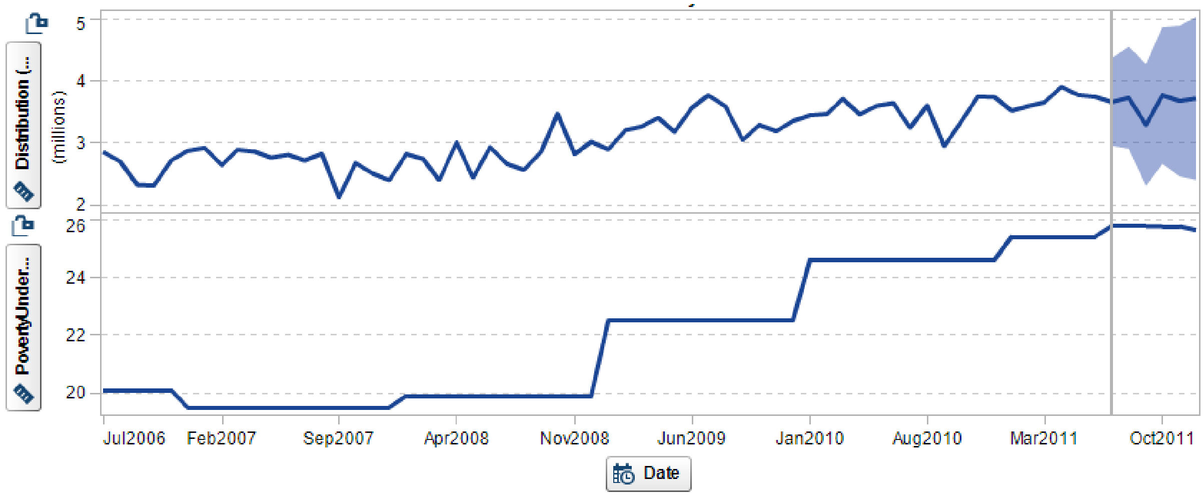

Figure 3: Forecast of food distribution with under-18 poverty introduced

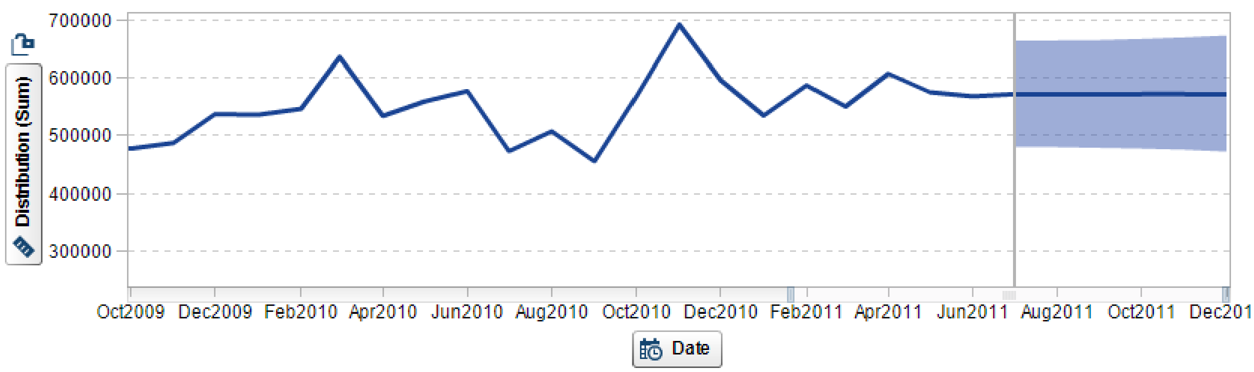

Figure 4: Forecast of food received by wake county (in pounds) – before introducing underlying factors

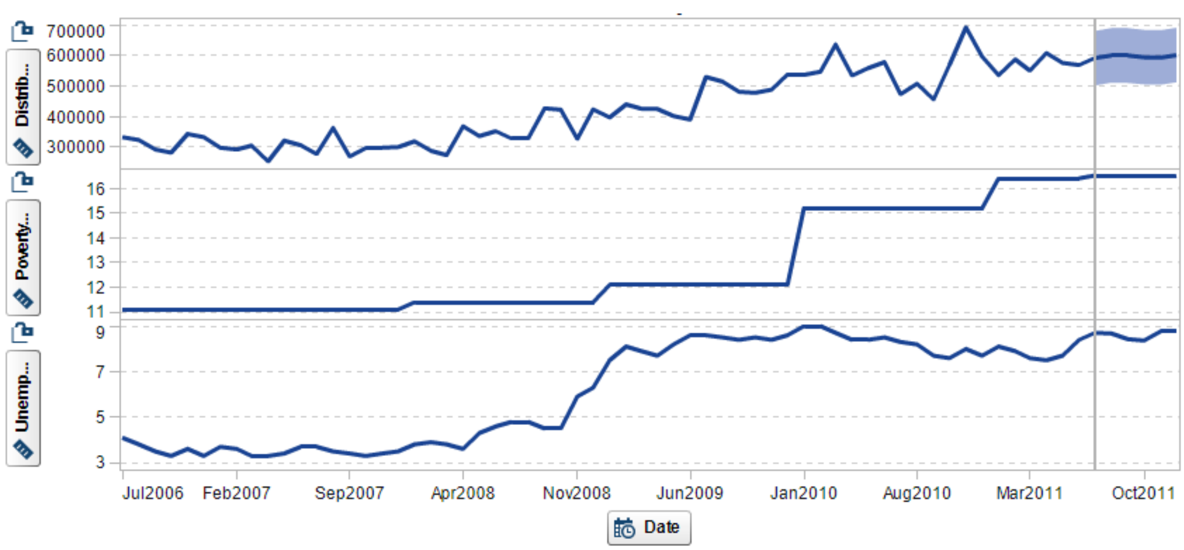

Figure 5: Forecast of food received by wake county – unemployment rate and under-18 poverty introduced

4. Discussion and Conclusion

In this research, we applied SAS VA to study the effect of underlying factors on food distribution forecasts. Results indicated that the introduction of the underlying factors affected the time series model selected by the system. Interestingly, under-18 poverty was identified as the significant contributor to the forecast for the entire food bank while both unemployment rate and under-18 poverty were determined to be the contributors to the forecasts for the county studied. Since there are variations among counties served by the food bank, great caution needs to be taken when we build times series models. Fortunately, VA tools such as SAS provide a quick and easy way for building the models and allow a food bank operations manager without statistics background to interact with the data and gain insights from the forecasts.

It needs to be pointed out that we selected several socio-economic factors based on our research as well as the availability of the data in this research. The food bank operations managers can identify more appropriate underlying factors based on their experience and therefore may improve the accuracy of the forecast.

Our next step is to perform scenario analysis and goal seeking. We will use SAS VA to visualize the changes made to the forecast by changing values of the significant underlying factors. In addition, we will set the forecast values to investigate changes to the underlying factors to achieve these values. The forecast values will be manually set to the gross weight of food distributed from the test data in the next period.

In summary, visual analytics provides a quick and easy way for operations managers to take advantage of the analytical power and yet apply their empirical experience to the forecast. This has the potential to improve the effectiveness and the efficiency of the food bank operations.

Acknowledgements

The authors would like to thank Charlie Hale and Earline Middleton of the Food Bank of Central and Eastern North Carolina (FBCENC) for their help and support.

This project was partially supported by NSF NRT: Improving Strategies for Hunger Relief and Food Security using Computational Data Science (Award No. DGE-1735258) and NSF PFI: BIC: Flexible, Equitable, Efficient and Effective Distribution (FEEED) (Award No. 1718671).

References

1. Coleman-Jensen, A., Gregory, C., Singh, A., 2013, "Household food security in the United States in 2013," USDA-ERS Economic Research Report (173).
2. Campbell, C.C., 1991. "Food insecurity: a nutritional outcome or a predictor variable?" Journal of Nutrition.121(3),408–415.
3. Coleman-Jensen, A., Rabbitt, M., Gregory, C., and Singh, A., 2018. "Household Food Security in the United States in 2017," ERR-256, U.S. Department of Agriculture, Economic Research Service (ii).
4. Food Bank of Central and Eastern North Carolina, 2018, "Hunger Studies, "Last retrieved January 16 2018. http://www.foodbankcenc.org/site/PageServer?pagename=hunger_studies
5. Delpish R., Jiang S., Davis L., Odubela K., 2018, "A Visual Analytics Approach to Combat Confirmation Bias for a Local Food Bank," In: Boring R. (eds) Advances in Human Error, Reliability, Resilience, and Performance. AHFE 2018. Advances in Intelligent Systems and Computing, vol 778, pp13-23. Springer, Cham.
6. Davis, L., Jiang, S., Morgan, S., Nuamah, I., Terry, J., 2016. "Analysis and prediction of food donation behavior for a domestic hunger relief organization." International Journal of Production Economics**,** 182, 26-37.
7. Shumway, R. H., Stoffer, D. S., Olkin, I., Casella, G., Fienberg, S., 2000. "Time series analysis and its applications". New York: Springer- Verlag.
8. Pereira, A., 2004. "Performance of time-series methods in forecasting the demand for red blood cell transfusion. Transfusion". 44. 739-46.

9. Chen, C., 1997, "Robustness properties of some forecasting methods for seasonal time series: A Monte Carlo study." International Journal of Forecasting, 13, 269–280.
10. United States Census Bureau, 2019, "Quick Facts." Last accessed January 16 2019, https://www.census.gov/quickfacts/fact/dashboard/wakecountynorthcarolina,US/PST045218.
11. Gardner, E. S., and McKenzie, E., 1985. "Forecasting trends in time series." Management Science*,* 31(10), 1237–1246.
12. Hyndman, R. J. and Athanasopoulos, G., 2018, "Forecasting: Principles and Practice". Second Print Edition. OTexts, Melbourne, Australia.