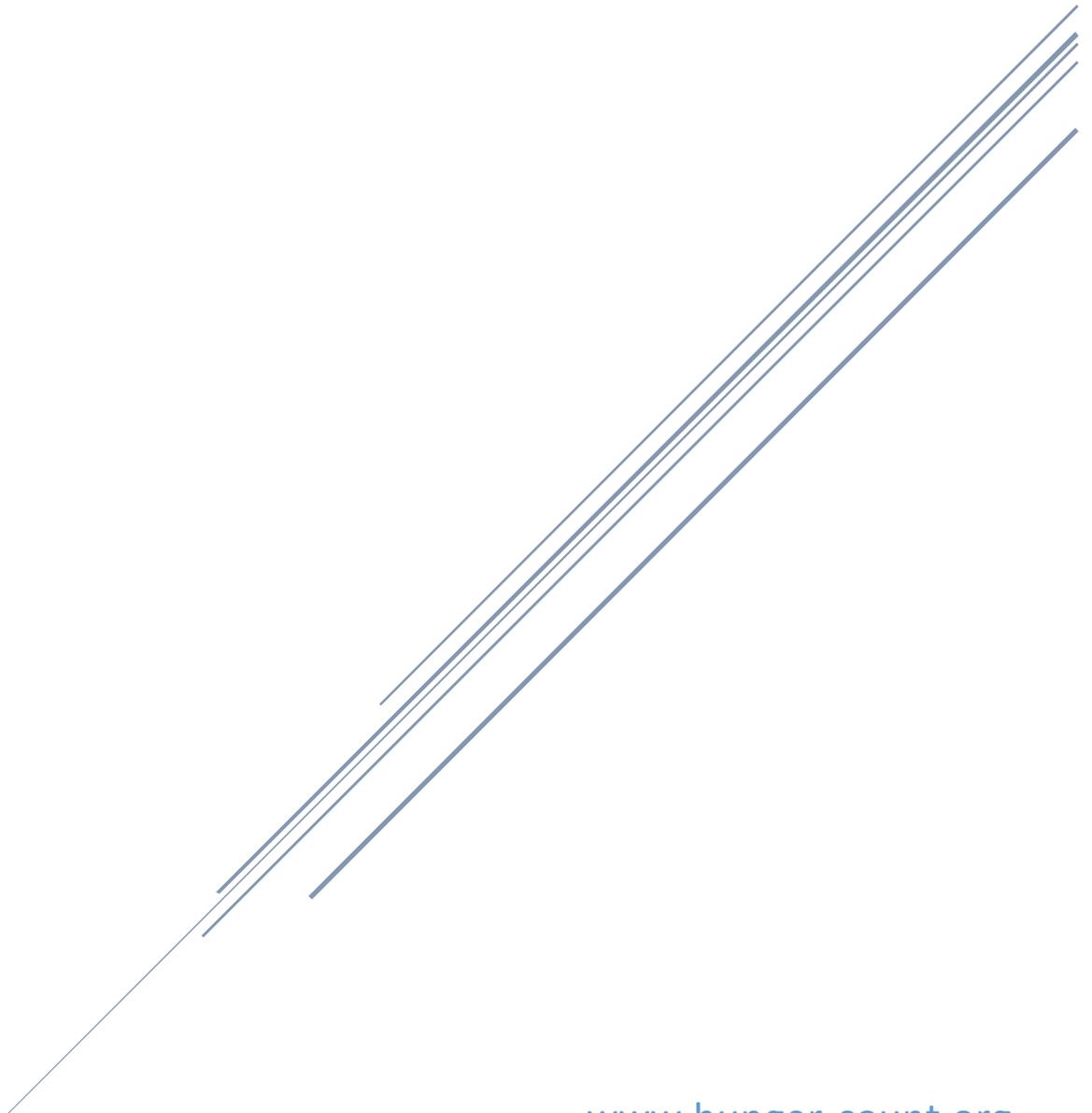


FRESNO HUNGER COUNT

Data Analysis Methodology



www.hunger-count.org

Written by Jacob Pence
Data Analyst

The initial analysis of quantitative data sought to derive basic information regarding food shortage within the Fresno Hunger Count survey area. The data was collected via interviews with residents within the survey area Census Tracts, which was then entered in to a web-based geographic information system (GIS) hosted by CSU Chico and ArcGISonline.com. Raw data could be downloaded from this online GIS database in an Excel format, based upon the type of data entered. This resulted in separate files for data entered for a “hungry” household, a “non-hungry”, a household that “refused”, and data entered from “rejected” survey forms.

In conjunction with CSU Chico Department of Geography and Planning, data was reviewed, and records with irreconcilable errors were removed (approximately 300 Hunger surveys, or 1.4%). This created a final set of records for data analysis to be done, containing 21,388 surveys denoting hunger conducted by FHC staff. This analysis focused specifically on the Census Tracts within the city limits of Fresno and Clovis, therefore ignoring any data points that existed outside this geographic delineation.

Data analysis was done using Microsoft Excel for three primary reasons. Firstly, when data was downloaded from the GIS database, or received from CSU Chico, it was already in the Excel format. Secondly, this data format is widely used, allowing a simpler way to exchange data files with staff of FHC without having to extract from a database system. This would not be ideal in larger applications, however, for an introductory study of small scope, such as this, it was effective. Thirdly, the types of analysis to be conducted, along with the limited scope of data, meant that additional database and analysis software would not be required.

Data required to derive useful information did not exist exclusively within the survey results. External data sets and information were used to supplement survey data to arrive at some conclusions, specifically in the following situations:

- (1) To find the dietary needs of individuals, which would allow responses to shortage to be extrapolated in to real world values, information from the United States Department of Agriculture, Department of Health and Human Services 2015 Dietary Guidelines Advisory Committee report and the World Health Organization report “Measuring intake of fruits and vegetables” was used. The USDA report was used to gauge the amount, by weight, of proteins, carbohydrates, and fats needed for an average adult to have a healthy 2,000 kcal per day diet. This report was not used to gauge fruit and vegetable intake because such intake was only estimated in volume, not weight, disallowing any accurate translation. The WHO report was then used instead, as it granted daily requirements of fruits and vegetables based upon weight. Dietary requirements for an adult may vary between such agencies, however, such variance is not thought to affect outcome validity.
- (2) The cost of food items needed was derived from multiple sources and aggregated to arrive at likely average costs for the area. The reason for this is two-fold. First; the survey was written with a focus on household’s common needs, specifying items, instead of surveying strictly for item groups. This limited the scope in aggregating prices, meaning a greater variance. As opposed to aggregating prices of *all*

protein/carbohydrate/fat items in the area, only those asked about on the survey were included. Second; due in part to this aspect of the survey, not all food items could be found in a single source. An alternative that would remove this issue would be to contact and survey suppliers in the area, however, this was deemed too inefficient for the project.

The following sources were then used to find average price of food items in the region: the USDA Economic Research Service, which gives the public access to a large amount of online data concerning the food industry. The Bureau of Labor Statistics Consumer Price Index database, which supplied the average price data for many items in the “Western Urban” market. Lastly, some items were not included in either of these sources. Olive oil and tortillas were instead given values per Phillip Erro based upon his interviews with local suppliers.

- (3) Data from the United States Census Bureau and by extension the American Community Survey was used throughout analysis. This includes, but is not limited to; estimating the amount of a census tract surveyors covered based on actual occupied homes, the number of children that likely exist in the area, average household size, the number of individuals below the poverty level, and fertility rates amongst financial levels.
- (4) Feeding America is a prominent organization in regards to researching and creating action against food insecurity within the United States. Access to Feeding America’s data allowed FHC to compare findings at a local level to estimates at a higher lever from a much larger organization with far more resources at their disposal. Community Food Bank handles the bulk of food distribution in the Fresno County area, and were also a valuable source of information regarding what foods are currently being distributed and how much.

Analysis of the survey data was focused primarily on questions four and five of the survey. These responses state the specific food items in each category that respondents needed, as well as which categories were required in which months. Question three was used to find at what times during the year insecure families were more common, while question six focused solely on homes with children and their insecurity. Questions one and two were meant simply to lead in to the others, and shift the respondent’s thoughts to their own experiences with food insecurity.

Support required by survey respondents was found by counting the amount of respondents that stated they were short of any item within a category for each month, and extending the cost to an average household based upon that value. The key obstacle in relating survey responses to the real world lies in question five and reflecting an insecurity in a month as an amount of time in that month. Baseline formulas give the total needs for a single person for an entire year, however, individuals do not need support for every meal every day of a year. The survey asked only for the month, not for how often in that month a family was insecure. However, a primary data analysis need was finding the cost and amount of food needed. The

find what percent of a month households were short of a category, the average percent of hunger respondents who stated they were short any item within a group during the year was found. That average was applied to the category from which it was derived to find an estimate of how much is actually needed.

The time of insecurity was found for commonality of months and days within a month by analyzing responses to question five. Monthly basis was derived by reference only to whether a respondent said yes for any shortage of any category within a month. There was very little variance highest and lowest month, so to better visualize the trend throughout the year, each month was plotted on a graph in relation to the mean number of responses.

The data representing the days during the month which households become insecure are skewed due to the construction of the survey. When asking for the date of insecurity, the survey form included the answers; “15th”, “20th”, “End”, and “Other”. The “End” value was entered as the 28th during data entry due to limitations of the database structure. Therefore, values were heavily skewed towards these three dates. This lessens the impact of analysis of such data, however, useful information can still be gleaned by separating a month into weeks; 1st to 7th, 8th to 15th, 16th to 22nd, and 23rd to 31st. Additionally, the values of 15, 20, and 28 were adjusted by taking only a small amount, 5%, in to account when tallying totals per week. This is because these values commonly had responses 40 to 50 times higher than non-specified dates. By making these adjustments, the shift of insecurity during the month can be easily visualized.

The survey and data entry method did not contain a way to accurately denote households which did not experience food insecurity and also had children. Data regarding small children within the home on hunger surveys was found via question six. Any answer to this shows that children are present, while answering question seven as well shows that there is also time the children are not food secure. To estimate the quality of this aspect of the survey, data from the Current Population Report regarding fertility of women in the U.S. was used, which allowed estimates for birth rate based upon income levels as compared to the poverty line.

Census data was used to estimate the number of homes with small children within the limits of the cities. A strict count is not available within the U.S. census or the American Community Survey, however, an estimated percentage of homes with children within specific age ranges does exist. Because no report gives information regarding the number of children per household, this estimate can only be seen as accurate assuming a max of one child within the age group per household.