



2015 SNAPshot Data Methodology

By David Thompson and JC Dwyer

SUMMARY:

2013 county-level data was provided by Feeding America's Map the Meal Gap project. See the methodology behind these numbers [here](#).

2014 zip code level data was compiled through several sources:

- Data for potentially SNAP-eligible (i.e. income-eligible) individuals was generated from American Community Survey (U.S. Census Bureau) 5-year population estimates of the number of persons at or below 125% of the Federal Poverty Line within the given zip code. These statistics were adjusted for county-level population growth to the current year using a multiplier derived from the Office of the State Demographer's Population Projections.
- The number and characteristics of SNAP recipients by zip code were provided to Feeding Texas by the Texas Health and Human Services Commission (HHSC) through a public information request.

All measures should be considered estimates and are meant for general understanding of the SNAP population in the state of Texas. The data provided is the best available, but anomalies may still occur. In particular, zip codes with very low populations, or which contain large student populations or group housing may present unusual results. As the data presented is only a proxy for the number of SNAP-eligible individuals in an area, this data should not be taken as an exact measure. Rather, it should be seen as a general guide for characterizing the local population receiving SNAP and improving SNAP outreach services to areas most in need.

The datasets powering the SNAPshot visualization tool are publicly available [here](#).

STEP-BY-STEP METHODOLOGY & NOTES:

(raw data available [here](#)).

1. Tabulating Total SNAP Receipts by Zip Code for the Year: HHSC spreadsheets for each month of 2014 were imported into Stata and saved as .dta files. We added an identifier variable for the appropriate month to each, then appended one upon the other until we had all 12 months worth of data into a single database. We then separately collapsed our data set to find monthly average receipts and to find total 2014 receipts and saved. This was later appended with the current eligible population data (below).

2. Projecting Current Eligible Population, Pt. 1 - Downloading Starting Population Estimates by Zip Code: Census 5 year estimates were taken by zip code, available at <http://factfinder.census.gov/>. Results were from the American Community Survey 5-year estimates and were found on form B17024. These results were downloaded into a .csv file and opened first in Excel. In Excel, we:

- Highlighted all numeric fields and formatted as a number
- Saved as an .xml or .csv file

We then opened up the saved table in Stata. See “Population Projections 2014” .do file for detailed instructions. In Stata, we:

- Deleted all “margin of error” variables
- Generated new variables using the “gen” command, adding results for each age group for 124% of FPL or less together.
- Dropped unneeded variables.
- Generated new variables for our three age categories of interest - 0-17, 18-64, and 65+. These are coded by adding other age group variables together.
- Saved.

3. Zip Codes missing from ACS: After mapping several of the Zip Codes that had been reported by HHSC but were not reported in the ACS data, we realized that many HHSC-reported zip codes pertained to post office boxes, business areas, and other nonresidential areas, as well as to Zip Code areas that had been reallocated from other preexisting Zip Code areas. During this process, we primarily consulted www.zip-codes.com to determine if the Zips pertain to a PO Box. By entering the Zip in question into the “Enter Zip Code” field to the left, you can pull up information on a particular Zip. We found that entering 78714, for example, returned info on that area, listing it as 0 population. This page also indicates that the Area Type is Z(PO Box Only). These Zip Codes then belonged to PO Boxes, indicating that the individuals listed by HHSC under these Zip Codes receive their SNAP mailings at their PO Box. This was not the case for Zip Codes 78599 and 75033. After searching

“78599 change usps” in Google we arrived at a news site for postal employees listing changes in Zip Code areas, which stated that some mailing addresses from 75034 had been reallocated to 78599 in 2011. It appears that the Census Department was unaware of this change when publishing ACS 2013 estimates. Therefore, we concluded that population estimates for both 75034 and 78599 were erroneous, and decided to modify them. To do this, we:

- Made an Excel file called “Used to Estimate Pop in Missing Zips”
- Found number of residential mailboxes in 75034 and 78599 from www.zip-codes.com.
- Made a multiplier for both zip codes by taking the number for 78599/(75034+78599)
- Multiplied the projected populations from our previous spreadsheets times these multipliers. These should be the current total current estimated populations.

4. Projecting Current Eligible Population, Pt. 2 - Estimating a Multiplier for Population Growth by County : After getting the per-zip code population estimate from the ACS, we need to estimate growth that has occurred since. The State Demographer provides population projections at <http://txsdc.utsa.edu/Data/TPEPP/Projections/Index.aspx>. Projected growth is available by county (at the lowest level), so we used county data to estimate growth in each zip code, i.e. we assumed that growth occurred uniformly in each zip code across the county. In our case, we were particularly interested in projections for the years 2011 and 2014. We assumed that the 2013 ACS data used above provides a population estimate most representative of the year 2011 (since 2011 is the median year in the 2009-2013 average provided by the 2013 ACS 5-Year Estimate.) Dividing the 2014 projected population by the 2011 projected population provides an estimate of the rate of population growth from 2011-2014.

Demography data is available via either a downloadable table on the website as “Population Projections for the State of Texas and Counties in one file” or via their Population Projections Tool at

<http://txsdc.utsa.edu/Data/TPEPP/Projections/Tool.aspx?fid=E2A8530641F4B01AB2C4733B66EE2DA>.

We cross-checked results from both the database and the projections tool and found that results were consistent with each other. Using the tool provides the most refined access to desired information, but requires some manipulation to avoid missed information and miscoding in Stata.

If you use the tool, make sure you follow these steps:

- Select all counties
- On the right hand side of the menu, do the following:
 1. Select 2000-2010 as the migration rate
 2. In the "Number of Years" box, enter "2011, 2014"
 3. In the "Age Group" box, enter "0-17,18-64,65-84,85+"
 - We were getting weird results when querying 65-120 and 65+, the Demographer's office suggested querying 65-84 and 84+ as above and combining results.
 4. In the "Sex" and "Race" boxes, click boxes for Total.
- Download .csv file
- Before opening in Stata, open in Excel. Select all fields with solely numbers in them and format as a number. This is important because we ran into coding problems in Stata when trying to convert these variables from string to numeric. This step is a good idea to implement when using the "Population Projections for the State of Texas and Counties in one file" table as well.
- Save and import to Stata. See the .do file "Population Projections 2014" for notes and commands used from this point on.

5. Preparing Master Working Stata Spreadsheet with Data for SNAPshot tool:

We collapsed our data set to find monthly average receipts saved. We then appended this with the current eligible population data. We generated and defined several variables from this information, discussed below. For detailed information, see the "Additional Stats Notes" .do file. The information we want in the SNAPshot spreadsheet includes the following variables. We will here walk through what needs to be done with each in Stata Spreadsheet:

averageBenefitperMeal: Generate variable = $\text{Avg_Dollars} / (\text{average number of days / month} * 3)$. The average number of days in a year is 365, divided by 12 and multiplied by 3 is 91.25.

constituentStory: Generate variable = ".". These may be filled in later by Story Project.

householdIncomeWithEarnedIncome: generate variable = $\text{ERND} / \text{FS_Cases}$

incomeEligible0To17, incomeEligible18To64, incomeEligible65Plus: Those eligible for food stamps.

incomeEligibleButNotReceiving0To17, incomeEligibleButNotReceiving18To64, incomeEligibleButNotReceiving65Plus: Those eligible but not receiving.

latitude, longitude, polygonCoords: Draw from previous .csv. Newer Zip coordinates such as 78599 and 75033 may be derived from GIS but we did not do this for 2014. We left them blank.

recipientEthnicity and recipientRace's: Those receiving SNAP from HHSC reports.

recipients0To17, recipients18To64, recipients65Plus: Self-explanatory.
totalIncomeEligibleButNotReceiving: Self-explanatory.
totalIncomeEligibleIndividuals: Self-explanatory.
totalSnapRecipients: Self-explanatory.

After you get your Stata spreadsheet headers in this format, export to .csv and you are done.