

FOOD DESERTS BY ZIPCODE

DATA SOURCES

How Big is the Problem in Three Largest U.S. Cities?



Los Angeles



Chicago



New York

Data Sources for Food Deserts by Zip Code

- There are three sets of data needed for this analysis
 - Income by Zip Code
 - Zip Codes in target cities
 - Grocery Stores near a Zip Code

The next pages explain these data sources and their use.

Data Sources – Income by Zip Code

- For income by zip code, this analysis will be using the 2017 income tax file available at <https://www.irs.gov/statistics/soi-tax-stats-individual-income-tax-statistics-2017-zip-code-data-soi>. The IRS data contains information by 6 major income groups for each zip code and indicates the individual line item information on U.S. Income Tax forms 1040, 1040A, 1040EZ. The information will be used to create an average income by zipcode for correlation and comparative purposes.
- The Data Dictionary for this file can be downloaded [here](#). (This will open in a Word instance.) Some rows from the Data Dictionary are below:

VARIABLE NAME	DESCRIPTION	VALUE/LINE REFERENCE	Type
STATEFIPS	The State Federal Information Processing System (FIPS) code	01-56	Char
STATE	The State associated with the ZIP code	Two-digit State abbreviation code	Char
ZIPCODE	5-digit Zip code		Char
AGI_STUB	Size of adjusted gross income	1 = \$1 under \$25,000 2 = \$25,000 under \$50,000 3 = \$50,000 under \$75,000 4 = \$75,000 under \$100,000 5 = \$100,000 under \$200,000 6 = \$200,000 or more	Num
N1	Number of returns		Num
MARS1	Number of single returns	Filing status is single	Num
MARS2	Number of joint returns	Filing status is married filing jointly	Num

Data Sources – Target City Zip Codes

- To retrieve zip codes for each city, this analysis will be using information from for www.zip-codes.com the three cities using the city names as keys for the data. The data will be scraped from the web site and refined as needed.

Screen shot of
Los Angeles Zip Code
data

The screenshot shows the ZIP-CODES.com website interface. At the top, there is a navigation bar with links: Home, Products, Learn About ZIP Codes, Find a Post Office, Search, Contact, FAQs, and Account Login. Below this is a header section with the ZIP-CODES.com logo and four database options: ZIP Code Database (US), Postal Code Database (Canada), ZIP Code Boundaries (US), and ZIP+4 Database (US). A green banner displays the phone number 1-800-425-1169 and a link to the Zip Code Database List.

On the left, the 'Our Products' section lists various services including US ZIP Code Database, US ZIP+4 Database, 2010 Census Database, Canadian Postal Codes, ZIP Code Boundary Data, ZIP Code API (marked as new), Area Code Database, Historical Zip Codes, FREE Radius Application, FREE Store Locator, FREE Mobile App, and NPA NXX Database. Below this is a 'Find ZIP+4 of Address' section with input fields for 'Address (U.S.)' (containing 'LOS ANGELES') and 'State Abbr', and a 'Find ZIP+4' button.

The main content area features a 'Free ZIP Code Finder' section with a search bar and a 'Questions? 1-800-425-1169' link. Below the search bar are tabs for 'Advanced Search', 'ZIP Code', 'ZIP+4 of Address', 'ZIP of City', 'Radius Finder', 'Distance Calculator', and 'Locate Post Office'. The 'ZIP of City' tab is selected, displaying a table of zip codes for Los Angeles.

Page 2 of 3

Zip Code	Area Code	City	County	State	Country
90045	323	LOS ANGELES	LOS ANGELES	CA	US
90046	323	LOS ANGELES	LOS ANGELES	CA	US
90047	323/213	LOS ANGELES	LOS ANGELES	CA	US
90048	323/310	LOS ANGELES	LOS ANGELES	CA	US
90049	818/310	LOS ANGELES	LOS ANGELES	CA	US
90050	424/213/323/818/626/747	LOS ANGELES	LOS ANGELES	CA	US
90051	424/213/323/818/626/747	LOS ANGELES	LOS ANGELES	CA	US
90052	424/213/323/818/626/747	LOS ANGELES	LOS ANGELES	CA	US
90053	424/213/323/818/626/747	LOS ANGELES	LOS ANGELES	CA	US
90054	424/213/323/818/626/747	LOS ANGELES	LOS ANGELES	CA	US
90055	424/213/323/818/626/747	LOS ANGELES	LOS ANGELES	CA	US
90056	323	LOS ANGELES	LOS ANGELES	CA	US
90057	213	LOS ANGELES	LOS ANGELES	CA	US

Data Sources – Grocery Stores by Zip Code

- To identify grocery stores by Zip code, FourSquare data will be used.
- The following *categoryId* keys will be used to identify the grocery stores:
 - grocery_store = '4bf58dd8d48988d118951735'
 - organic_grocery = '52f2ab2ebcbc57f1066b8b45'
 - supermarket = '52f2ab2ebcbc57f1066b8b46'
- The range for the grocery stores will be 804 meters, approximately ½ mile. This is in line with the USDA's definition of reasonable grocery store access.
- Additional information about venue categories within FourSquare can be found [here](#).

The logo for FourSquare, featuring the word "FOURSQUARE" in a bold, blue, sans-serif font. The letters are all in uppercase. The logo is set against a white rectangular background.