# Project 1: Food deserts in the US

| **Objective:** | Replicate previous study with new data |
| --- | --- |
| **Description:** | Review status of food deserts in 5 US locations, looking at supermarket density and census data. |
| **Question 1:** | Has there been improvement at our 5 locations on availability of supermarkets?  Napa, CA - Rural (Napa County) Bronx, NY - Suburban/Urban (Bronx County) San Francisco (Bay Area) - Urban (San Francisco County) Victorville, CA - Suburban (San Bernardino County) |
| **Question 2:** | Is there a correlation between poverty rate vs food deserts? Different variables: |
| **Question 3:** | Is there a correlation between population and food deserts? |
| **Question 4:** | Has there been improvement on availability of supermarkets in the US overall? |
| **Conclusions** | Availability of food at our 5 cities, US overall, improvement since 2010 (original study), next steps |

## What is a Food Desert?

The USDA defines living in a food desert as:

living more than one mile from a supermarket in urban/suburban areas, and more than 10 miles from a supermarket in rural areas

This data is pulled from the [**Food Access Research Atlas**](https://www.ers.usda.gov/data-products/food-access-research-atlas/documentation/), and contains information on supermarket access at various distances. This data measures access by the Census-Tract, and as such provides a fairly granular overview. Additionally, what I thought was interesting is that it combined Food Access data with other fields such as age, race, rural or urban, and income.

## Previous study

<https://www.kaggle.com/code/kerneler/starter-food-deserts-in-the-u-s-3f49ae84-7/input>

<https://www.ers.usda.gov/data-products/food-access-research-atlas/documentation/>

## Current datasets

#### Geoapify (Supermarket Density)

https://api.geoapify.com/v2/places

#### US Census (Demographic data of locations)

<https://www.census.gov/data/developers/guidance/api-user-guide.html>

#### USDA Economic Research Service (Supermarket Access)

[https://www.ers.usda.gov/data-products/food-access-research-atlas/documentation](https://www.ers.usda.gov/data-products/food-access-research-atlas/documentation/)

## Next Steps 1/10/2024

* Review previous study
* Review current datasets
* Extra team session planned for Sunday Jan 14

References:

Example from class: study of banking deserts can inform study of food deserts 01-Sessions/06-Python-APIs/3/Activities/09-Evr\_Banking\_Deserts/Solved/banking\_deserts\_solution.ipynb

Example from class: census data  
06-Python-APIs/3/Activities/08-Ins\_Census/Solved/census\_demo\_solution.ipynb

Census labels  
https://gist.github.com/afhaque/60558290d6efd892351c4b64e5c01e9b

Small Area Income and Poverty Estimates: State and County  
<http://api.census.gov/data/timeseries/poverty/saipe>

## >Join supermarket data with census data: merge on zipcode DONE

Create dataframe that include supermarket count (for loop to capture json results)

Clean dataframe (dropna, fillna)

Add any additional calculated columns

Sort data

Plot data

Question:

## >Group data by counties (based on zipcode)

## >Map supermarket availability

## >Highlight averages for our 5 cities compared to national averages (mean, median, mode)

>Calculate r-squared value

>Plot supermarket availability against population (scatter plot with regression line)

## 

## Archive

Fast food restaurants across US

<https://www.kaggle.com/datasets/khushishahh/fast-food-restaurants-across-us/data>

Restaurant data by neighborhood

<https://apidocs.geoapify.com/docs/place-details>

Restaurant & consumer data according to consumer preferences  
<https://archive.ics.uci.edu/dataset/232/restaurant+consumer+data>

Characteristics and Influential Factors of Food Deserts USDA

<https://www.ers.usda.gov/webdocs/publications/45014/30940_err140.pdf>

Mapping Food Deserts in the United States USDA

<https://www.ers.usda.gov/amber-waves/2011/december/data-feature-mapping-food-deserts-in-the-u-s/>

From Alex: includes latitude and longitude data

<https://www.kaggle.com/datasets/khushishahh/fast-food-restaurants-across-us/data>

1. Establish Question -   
   Comparison of fast food restaurants between different areas, rural, urban, suburban???  
   Supermarket vs fast food restaurants - availability between the two in different areas???
2. Establish Dataset - Geoapify (if time permits) - to find restaurants & supermarkets using categorization to pull the API

Kaggle - CSV file with data about either the restaurant data by neighborhood OR food deserts in the US in general.

Sharon: API connection to <https://api.geoapify.com/v2/places>, and dataset by 5 cities, by category (restaurant, supermarket)

Nathan: Kaggle data analysis concentration of restaurants–can it be filtered by our 5 cities?

Mary: Reviewing datasources we collected, presentation/write-up

Aman: Reviewing datasources for trends

(Alex: Slack, Github, review definition of food deserts) ?

Kaggle - Food desert

Project Stages

* Project ideation
* Data fetching/API integration
* Data analysis
* Testing
* Creating documentation
* Creating the presentation

Food Desert CSV data Summary   
  
Explain relevant csv column titles   
Census tract  
County   
LA1and10   
POP2010

PovertyRate

Using the food\_access\_research\_atlas data from 2010, we filtered by the column LA1and10 which showed whether or not a census tract was a food desert or not.   
  
Upon initial analysis, we saw that Napa County, San Francisco County, and Bronx County were all relatively similar in their lack of food deserts. However, the outlier was San Bernandino.   
  
We then proceeded to adjust our code so that it pulled the population of the respective counties we were looking at while also still denoting if it was a food desert or not.   
  
We decided to focus on the Urban/non-urban factor to see if that also correlated to being a food desert. However, we found that there was no direct correlation between food desert and these areas being urban/non-urban.  
  
  
  
We decided to focus heavier on San Bernardino county because of this in order to determine