**Use either Power BI OR R-Script (R-markdown files) OR Excel for the below analysis.**

**Analyze food access and mortality in US.**

Use this data set:



We're interested in looking at the relationships between food access, mortality, and politics.

To do this, we look at data from three different sources:

- The USDA's [Food Environment Atlas](https://www.ers.usda.gov/data-products/food-environment-atlas/documentation/)

- The CDC's ["Compressed Mortality File 1999-2015 Series 20 No. 2U, 2016"](http://wonder.cdc.gov/cmf-icd10.html)

- 2016 election results (found all over the internet)

Each row in the dataset is a US county. The main outcome we care about is `mortality\_rate`, or the number of deaths per 100,000 people in a county between 2013-2015. Other interesting variables in the dataset include:

- `pct\_low\_access\_pop`: Percent of the county's population with low access to food

- `pct\_children\_low\_access`: Percent of the county's children with low access to food

- `grocery\_stores\_per\_1000`: Number of grocery stores in a county (per 1,000 residents)

- `snap\_stores\_per\_1000`: Number of stores that accept SNAP (food stamps) in a county (per 1,000 residents)

- `fastfood\_per\_1000`: Number of fast food stores in a county (per 1,000 residents)

- `per\_dem\_2012`: Percent of the county that voted for Obama in 2012

- `per\_dem\_2016`: Percent of the county that voted for Clinton in 2016

You have the option to use Power BI or R or Excel to complete this exercise.

If using Power BI submit a complete pbix file. If using R submit an executed R-markdown file that has your code, comments and any plots. For excel, send back the excel file.

Please ensure any analysis has commentary, and any plots have axes and key insights called out.

Perform the following:

1. Analyze the data set for completeness, and integrity, missing data etc.

1. Carry out exploratory data analysis.
   1. Try to investigate the following:
      1. How related are mortality rate and access to food?
      2. How related are mortality rate and the prevalence of fast food restaurants?
      3. How related are mortality rate and the prevalence of SNAP stores per 1,000 residents?
      4. How related are mortality rate and the percent of the county that voted for Democrats in 2016?
      5. Top 5 states that have the widest variance in "Percent of the county's population with low access to food" for county's that fall within that state?
      6. Top 3 states that have in absolute terms, reduced most the population with low access to food across the election years 2012 and 2016? Columns "low\_access\_pop" and "low\_access\_change" have relevant data.
   2. Any other interesting insights that you can discover from the data?
2. Additional insights: Try to explore the questions below:
   1. Does access to food predict mortality?
   2. Do more SNAP stores per person predict mortality?
   3. Do election results and access to food and SNAP stores predict mortality?