Group G: Moises Escobar, Kexu Duan, Justin Ko, Yiwei Li

Professor Brambor

G5063: Data Visualization

03/01/2023

Group G Project Process Book: Health and Nutrition Patterns in the United States

Summary

Our group aims to use the skills developed from this data visualization course to focus on an array of narratives pertaining to health and nutrition data in the United States. We intend to explore, analyze, and visually present the following:

1. Explore the current state and relationship of health and nutrition in the United States (comparisons by state, race, etc)

2. Analyze how the trend has changed over time (time series)

3. Identify the challenges to fixing malnutrition and obesity in the United States

4. Display the potential health outcomes for increased spending in the healthcare and education sector in the United States

5. Recommendations to improve longevity and well-being for people in the United States

Sections

1. **The current problem** – Obesity epidemic spreading across the U.S

    a. Inter-state comparisons, i.e., 50 states

    b. Inter-race comparisons, i.e., Different races in the US

    c. International comparisons, i.e, US vs developed countries in G7, or US vs other countries in G20

    d. Word cloud of people's favorite food / high school meals

2. **Time-series**

    a. How have people's health deteriorated (or improved in some respects) in the US over the years – line graphs, heat maps (interactive)

3. Why is it so challenging to fix? What are the variables involved?

    a. Fast food location geospatial maps

    b. People do not really know what to eat — the food pyramid and measure of healthy foods keep changing

    c. Ultra-processed food (UPF) consumption is very high due to its convenience and a large segment of the population does not know how it negatively affects their health

4. Attempts at improving the health of the general public

    i. Vegan / vegetarianism

    ii. Other health fads: paleo, intermittent fasting, etc

        1. Building graphs based on the number of Google searches

        2. Studies pertaining to weight loss as a result of new diet plans

5. Outcome (What are the potential solutions?)

    a. Increased healthcare spending in the US vs. The healthcare spending in other developed countries in G7

    b. Comparisons between average body weight (sex, race, state, nationality)

    c. Mortality rates

        i. Inequality: The lifespan of top income decile in the US is probably high

6. Education/Accessibility (location and prices)

    a. Where to put healthier food options (interactive maps)

    b. Subsidies to fast food (interactive graphs with ggplotly)

    c. Education - How it starts with you (Data table of foods' nutritional information)

Visualizations

1. Static images based on ggplot2

    a. Line graphs with percentages of Americans with different health conditions across time

    b. Line graphs comparing the United States to other G20 countries on mortality across time

    c. Mortality rates in the US by income decile

    d. Flipped axes bar graph showing data for a panel of different health conditions by state

    e. Flipped axes bar graph showing data for a panel of different health conditions by race

    f. Something around the rise of vegan / vegetarianism

        i. Has it plateaued?

    g. Facet wrap of knowledge quiz responses

    h. Facet wrap of diet responses

2. Maps using geospatial data

    a. Fast-food restaurant locations

    b. A map of the United States showing the percentage of fast food restaurants out of total restaurants in a given county or state (shade from blue to red by percentage)

    c. A panel of a map of the United States showing the percentage of a given health condition in a given county or state (shade from blue to red by percentage)

    d. Obesity statistics across time (1950, 1960, …, 2020)

3. Visualization of text analyses

      a.   Cities in America with the least healthy populations – the size of the text

            i.   Across different metrics

      b.   People's favorite food / served in public high schools

      c.   Panels with the most common fast food restaurant in a given state – size of text

      d.   Foods with the greatest amount of saturated fat / [unhealthy substance]– the size

          of the text

      e.   Which foods are mentioned most often in state newspapers or other examples

          from Lecture 8

4.   Network visualizations

      a.   [xx]

5.   Hosted, interactive display of some visualizations

      a.   Show the United States map getting plumper over time, i.e., inflating

      b.   Something around the growth of UPFs over time

             i.   It's a GIF and the food pops up in the year the product was released

      c.   GIF of how the food pyramid has changed over time with food icons moving up

          and down the levels

      d.   Data tables with foods and columns give you nutritional information, i.e., calories,

   saturated fat, trans fat, vitamins, etc.

Meeting Notes

1) March 09:

- Knick off

- Discuss the outline of website

The goal for the final output of each group project is a visually appealing, easy to use, informative, and (partially) interactive website. In the initial development of the project, feel free to keep individual results / graphs etc. in separate, self-contained web pages with visualizations (and maybe text, other info, some links etc.). Towards the end of the course, we will cover how to combine these separate pages to make a simple final project website and host it.

2) April 16:

1. The current problem (Tab #1)
    a. <mark style="background-color:yellow">Moises</mark> - Geospatial map showing obesity by state across time
        i. Perhaps make interactive so user can select the year
        ii. Data:
        iii. Geospatial map of the U.S. where major cities are written by obesity %, i.e., least healthy cities are written larger than smaller ones
    b. Moises - Histograms of responses to survey questions
        i. Knowledge quiz
        ii. Consumption habits
        iii. Data: Centiment survey
    c. <mark style="background-color:#00ff00">Moises</mark> - Histograms comparing health stats by race, income, education level
        i. Data: Centiment survey
    d. Moises / Leo - Pyramid plot of obesity by age group between male and female
        i. Data: Centiment survey
        ii. [BRFSS Prevalence & Trends Data: Explore by Topic | DPH | CDC](#)
        iii. [Chronic Disease Indicators: Explore by Indicator | DPH | CDC](#)
        iv. [Healthiest Counties in America Based on Obesity Trends | County Rankings (medicareadvantage.com)](#)
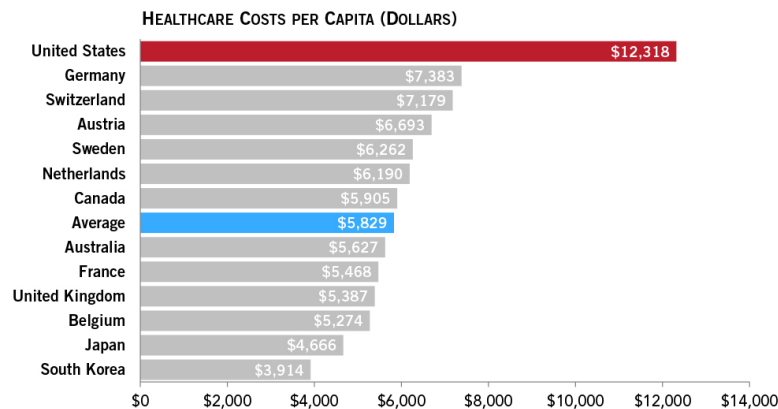   v. [Data & Statistics | Overweight & Obesity | CDC](#)

██ MALES ██ FEMALES
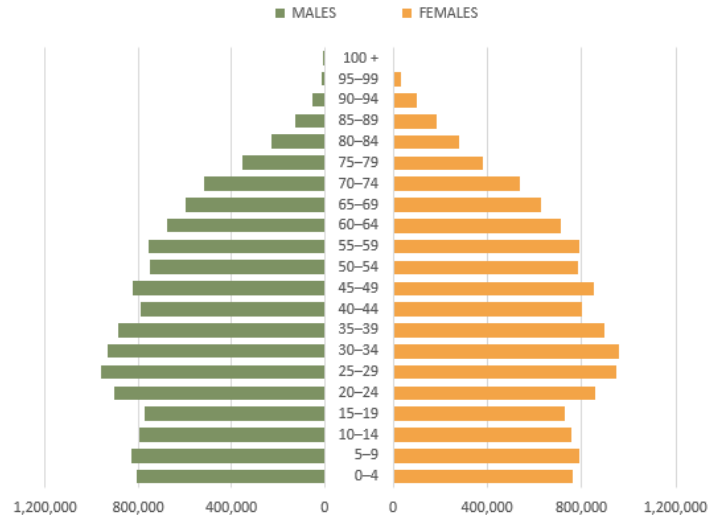
   vi. EXAMPLE:

 e. Kexu - Make a word cloud of people's favorite food in the U.S. or people's favorite restaurant
   i. Data: Fast Food Restaurants Across America | Kaggle
   ii. Interactive

2. Challenges to healthier diets (Tab #2)
 a. Leo / Justin - Interactive Geospatial map of fast food locations across the United States or in a given state / city (e.g., NYC)
   i. Data: Fast Food Restaurants Across America | Kaggle
 b. [Map of the US with the most popular fast food chain by state]
 c. Justin - Histogram (flipped axis) of healthcare spending in the U.S. vs other countries

PETER G. PETERSON FOUNDATION

U.S. per capita healthcare spending is over twice the average of other wealthy countries

HEALTHCARE COSTS PER CAPITA (DOLLARS)

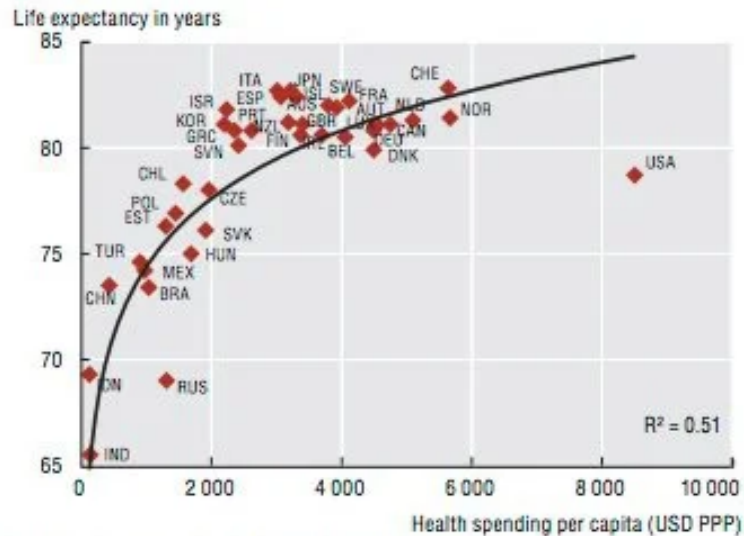| Country	Cost
United States	$12,318
Germany	$7,383
Switzerland	$7,179
Austria	$6,693
Sweden	$6,262
Netherlands	$6,190
Canada	$5,905
Average	$5,829
Australia	$5,627
France	$5,468
United Kingdom	$5,387
Belgium	$5,274
Japan	$4,666
South Korea	$3,914

SOURCE: Organisation for Economic Co-operation and Development, *OECD Health Statistics 2022*, July 2022.
NOTES: Data are latest available, which was 2019, 2020, or 2021. Average does not include the United States. The five countries with the largest economies and those with both an above median GDP and GDP per capita, relative to all OECD countries, were included. Chart uses purchasing power parities to convert data into U.S. dollars.
© 2022 Peter G. Peterson Foundation   PGPF.ORG

   i.
 d. Kexu - Scatterplot of healthcare spend on x avis vs healthcare outcomes (e.g., life expectancy)

Life expectancy in years

Source: OECD Health Statistics 2013, http://dx.doi.org/10.1787/health-data-en;
World Bank for non-OECD countries.

StatLink ▀▄▄ http://dx.doi.org/10.1787/888932916040

     i.

e. Kexu - Building a food pyramid with ggplot2
- i. Maybe show what the FDA recommends and then have labels calling out the "errors" in these recommendations
- ii. Maybe a GIF of how the food pyramid has changed over time with food icons moving up and down the levels
  - 1. Or make 5 food pyramids, one for 1980, 1990, 2000, 2010, 2020

f. Leo - Some visualization for school funding towards quality food in public schools
- i. How it changed over time
- ii. Michelle Obama's work in this space ([Healthy, Hunger-Free Kids Act of 2010 - Wikipedia](#))

g. Justin - Education - How it starts with you (Foods' nutritional information)
- i. Dataframe: searchable data table that shows the most popular food, the types of food, processed vs. unprocessed, and its respective calories
- ii. Data:https://www.kaggle.com/datasets/niharika41298/nutrition-details-for-most-common-foods
- iii. Potential solutions (Tab #3)
  - 1. Leo - Bar Chart with fitted line / Line chart: highlight the trend or rise of health fads and dietary options: vegan/vegetarianism, paleo, intermittent fasting, etc

There will also need to be quite a bit of text written around our visualizations to describe our thoughts / to make the project flow better. Just flagging it now so we can think about assigning that responsibility.

3) April 23:

1. The current problem
   a. Heat map of obesity rates showing trends and regional disparities - <mark>Moises</mark>
      i. How to make it animated so that the same map updates every 5 seconds for the next year
         1. gganimate (add a thicker border along regions to highlight regional disparities)
         2. Check Prof lectures for code
   b. BMI across time - <mark>Moises</mark>
      i. Find data and then try to make animated like above – see gganimate gapminder example
      ii. BMI calculator - <mark>Leo</mark>
         1. Use shiny to add calcultor to website
         2. drinkr/server.R at master · rasmusab/drinkr · GitHub
2. Challenges to healthier diets
   a. Fast food accessibility - <mark>Justin</mark>
      i. Focus just on nyc and aggregate number of restaurants (hopefully we'll see that there is a higher density of fast food restaurants in poorer areas of the city)
      ii. Try adding logos for restaurant locations
      iii. Think about how to add a useful legend
         1. Color coding the restaurants on the map
         2. Maybe the count / percentage?
   b. Favorite fast food restaurant word cloud - <mark>Kexu</mark>
      i. Check all types of restaurant data
      ii. Add percent for restaurant count
      iii. Try to add logos instead of restaurant names
         1. If this is too difficult, then just add a few logos around the word cloud for the top restaurants
   c. Top fast food brand presence by geographies
      i. focus on top 7 fast brand brands from word cloud only
      ii. split the country into 4 key regions
   ~~d. Food consumption expenditures - Kexu~~
      i. Line graph of expenditures on different food groups (fruits, vegetables, protein, snacks, grains, instant food, desserts, etc.) across time
         1. Think about whether you want proportion or a dollar amount on the y-axis
            a. If it's a dollar amount, it could be that certain foods are getting more expensive OR that people's consumption patterns are changing
   e. Life expectancy vs healthcare spend per capita - <mark>Kexu</mark>
      i. Animate by having data across years – ask Moises if he can figure it out above and use similar code
      ii. Add popup when you hover over a data point including flag emoji, exact data for life expectancy and healthcare spend per capita, and population

        iii.   Maybe label only the US (and perhaps a few more like China, United Kingdom, France, Mexico, Brazil, South Africa, Kenya, Sudan) not every other country

3. Potential solutions
    a. Different diets people follow across time - <mark style="background-color:green">Justin</mark>
        i. Try to find a better data set including vegan, vegetarian, paleo, keto, intermittent fasting, etc.
    b. Geospatial map showing vegan / vegetarian restaurants in top 3 healthiest and unhealthiest states - <mark style="background-color:yellow">Justin</mark>
        i. Moises to send Justin the names for 2016 - Done
        ii. Healthiest states: Colorado, Massachusetts, Washington DC (or Hawaii)
        iii. Least healthiest states: Mississippi, West Virginia, Arkansas
    c. Education
        i. Healthcare spent on healthier foods in public schools - <mark style="background-color:red">Moises / Leo</mark>
            1. Try to find data - maybe think about line graph to show changes across time
            2. Research Michelle Obama's work and see if she focused on specific areas or schools
        ii. ~~Food pyramid -~~ <mark style="background-color:red">~~Leo / Kexu~~</mark>
            1. ~~Try to find different food pyramids across time in the US or across cultures (US vs China vs Europe etc)~~
        iii. Data frame - <mark style="background-color:green">Justin</mark>
            1. Make title larger font
            2. Figure out what t means in data table or fix if it's a mistake
            3. Consider coloring food name background by Category (i.e., Dairy products, etc)
4. Putting everything together on a shiny app - <mark style="background-color:green">Moises</mark>

5. Think about code book to submit with project

NEXT STEPS (04/30)
1. All to include rmd files in shared repo
2. Justin to produce geospatial map showing vegan / vegetarian restaurants in top 3 healthiest and unhealthiest states
3. Moises and Dora to make missing edits to graphs
4. Leo to keep an eye out for next steps on his graphs

4) May 02:

1. Put final graphs on single file
    i. Heat maps of obesity
    ii. BMI
    iii. BMI calculator
    b. Knit to HTML
    c. This is the document we will present from

2. Write script
3. Record video
   a. Flip to shiny app after showing visuals and explain that next steps are to put everything on shiny app and edit according to classmate comments

4. Theme? Use ggthemes package (maybe theme_stata())
- https://mran.microsoft.com/snapshot/2017-02-04/web/packages/ggthemes/vignettes/ggthemes.html
- https://r-charts.com/ggplot2/themes/

Example visualizations
https://github.com/QMSS-G5063-2023/course_content/blob/main/Lectures/Week03/week03_lecture.md