**Dataset Background:**

The dataset we chose to analyze is from the Behavioral Risk Factor Surveillance System.

The BRFSS is a collaborative project between all states in the U.S. and the CDC.

It is the nation’s premier system of health-related telephone surveys that collects data about U.S. residents regarding their health-related risk behaviors, chronic health conditions, and use of preventive services.

Established in 1984, the system completes more than 400,000 interviews of adults each year, making it the largest continuously conducted health survey system in the world.

The results of the surveys are combined and made available to the public through the CDC’s website.

This data has become a powerful tool to help the nation and individual states establish and track local health objectives, plan health programs, implement disease prevention and health promotion activities, and monitor trends.

Nearly two thirds of states use BRFSS data to support health-related legislative efforts.

*Source:*

*-https://www.cdc.gov/brfss/about/index.htm*

**Introduction to Our Analysis**

Using this dataset, we set out to answer the following questions:

* What percent of adults aged 18 years and older have obesity?
* What percent of adults achieve at least 150 minutes a week of moderate exercise or 75 minutes a week of vigorous exercise?

The CDC defines obesity as an adult who has a BMI of 30 or higher. For this dataset, the BMI was calculated using self-reported height and weight from survey respondents.

While the dataset included responses to multiple levels of physical activity, we chose to look at the percent of adults who achieve at least 150 minutes of moderate exercise as this is what is recommended for adults by the U.S. Department of Health and Human Services according to the current *Physical Activity Guidelines for Americans*.

This version of the BRFSS dataset is a summation of survey responses from 2011 through 2021 for all states and 3 territories in the U.S. The responses are categorized by:

* Age
* Education
* Gender
* Income
* Race/Ethnicity

We analyzed these two questions by Income, Time, Geography, and Sex.

**Income:**

Data was aggregated to a national level and all other stratifications were summed. Income was analyzed in the following categories:

* Less than $15,000
* $15,000 – $24,999
* $25,000 – $34,999
* $35,000 – $49,999
* $50,000 – $74,999
* $75,000 or greater

The data showed an inverse relationship between the rate of obesity and income with an r value of 0.95. The data showed a positive relationship between exercise and income with an r value of 0.99.

The data confirms that individuals who have higher levels of income tend to engage in exercise more often and have lower levels of obesity.

Significant data points for comparison:

* % who Exercise:
  + Less than $15,000 – 41%
  + $75,000 or greater – 60%
* % who are obese:
  + Less than $15,000 – 35%
  + $75,000 or greater – 26%

**Time**

The data frame by time was filtered to include responses from all locations and for all demographic categories before being sorted by year. A new data frame was then created in response to each question, and line graphs were created as visuals.

Obesity over time:

* From 2011 to 2021 the rate of obesity in survey respondents from across the nation rose from 27.4 to 33% for a total increase of 18.5%

Physical Activity over time:

* This question was only included in the survey every other year from 2011 through 2019. A 2.8% decrease was initially seen from 2011 to 2013. However, over the entire eight-year range the percentage of respondents achieving at least 150 minutes of moderate exercise each week remained almost unchanged with a less than 1% decrease.

Time Conclusion:

* Interestingly, obesity has climbed over the past decade while the percentage of respondents achieving recommended levels of physical activity has remained steady. This suggests that physical activity is not the sole determinant of obesity and highlights the complexity of the disease.
* It may also suggest that the effectiveness of national guidelines and health policies related to physical activity levels of adults stagnated during this period.

**Geography (State):**

The bar graph in the slide deck shows the relationship between location, and the percent of adults with obesity.

* As shown on the graph, the states with the highest levels of obesity in adults are, Alabama, Arkansas, Louisiana, Mississippi, Oklahoma all within the 34-36%
* The lowest percentage of adults with obesity were found in in Colorado, Washington DC, Hawaii, and Massachusetts all within the 21-23% range

It is important to note that there is an obvious direct correlation between adult obesity and income. The states with more adult obesity also have more adults in the low income bracket. States with a lower percentage of obese adults have higher income. The income section confirms this very statement.

Physical Activity levels graph:

* Switching to the percent of adults meeting the weekly physical activity guidelines
* There is no surprise that the states with a higher percent of adults meeting physical activity guidelines are almost the exact same states with lower obesity rates.
* The states with a lower percentage of adults who meet these guidelines, also directly correlate to the obesity levels

\*THERE ARE A LOT OF DIFFERENT FACTORS THAT AFFECTED EACH STATES OUTCOME. A few are:

* Different policies- investments and resources that promote healthy lifestyles
* Poverty rates in each state, like we discussed

Next steps: Something to point out, some states have larger sample sizes. So a next step for me particularly, would be to dig into the sample size to understand if it's relative to the state population.

**Sex over time:**

* Activity levels are fairly consistent across time for both genders while obesity climbs for both men and women over time
* Slight yet steady increase in obesity over time
  + Even though the increases are fractions of percent values, since this data only spans a six year period, I would consider the increase to be a significant cause for concern, especially if obesity rates were to continue to climb at the same rate over a period of decades.
* Although women tend to be less active than men, activity levels are still roughly gender-equal
  + This tracks with findings that women typically have less leisure time than men, are less likely to be employed in highly physically demanding jobs, and may choose to spend their leisure time pursuing more sedentary hobbies (source: *Invisible Women,* Caroline Criado Perez, 2019). Since obesity levels appear to be unaffected by this exercise discrepancy between sexes, gender-specific intervention is likely unnecessary in order to improve health outcomes.

**Sex across location:**

* Once again, sex disaggregation reveals that exercise levels and obesity rates are more or less gender-equal across locations, with women being slightly less active than men overall, but not enough to result in significant discrepancies in health outcomes.
* There is a clear inverse relationship between activity level and obesity rates.
* Sex appears to have a negligible effect on obesity rates. States with poor health outcomes in this regard appear to fit at least one of these three criteria: poor, southern, non-state/territory.

The next step would be to compare the sex ratios in the sample sizes by state to the sex ratio of the entire state using census data. I made data tables to go along with all of my dataframes, and when I scanned everything with my bare eyes, the sex ratio of the sample sizes appeared to be about 50/50, but I would want to actually crunch the numbers before assuming that my data was representative of real life populations.

**Next Steps:**

Potential next steps for the project could be to:

* Use data to determine the effectiveness of current or previous health policies at national or local levels. Because the “income” category revealed the most stark and direct inequality levels, health policy researchers could investigate the states with the best overall health outcomes to see what policies or infrastructural changes could be implemented in lower-performing areas.
* Use data to determine what demographics should be prioritized by new health initiatives to equitably reduce obesity in the United States.
* Use predictive modeling to determine where resources should be allocated decades in advance in order to create a more robust national public health system.