

## Abstract:

I conducted this project to better understand the relationship between physical activity and obesity rates across U.S. states. Given the growing public health concern surrounding obesity, I was interested in examining whether differences in physical activity levels could help explain variation in obesity prevalence. By analyzing state-level data, I aimed to identify patterns that could highlight the importance of lifestyle behaviors in influencing health outcomes. Overall, the findings suggest that higher levels of physical activity are associated with lower obesity rates, reinforcing the idea that increasing physical activity may play a key role in improving population health.

## Data & Sources:

**Data:** This dataset consists of 20 observations chosen at random, with each observation representing an individual U.S. state. The data includes health, behavioral, and demographic characteristics used to examine relationships between physical activity and obesity rates across states.

**Method & Sources:** This project used a research method analyzing available data collected from reputable databases. Health-related variables, including obesity rates and physical activity levels, were obtained from the *Centers for Disease Control and Prevention (CDC)*. Demographic and socioeconomic variables, such as median income and educational attainment, were obtained from the *U.S. Census Bureau* and related federal datasets. Data regarding food access, including fast food restaurant density and fruit and vegetable consumption, were compiled from national public health and nutrition databases. The data represents a sample of U.S. states and are intended to reflect broad patterns across the United States. Because the data was collected from credible, government-based sources, they are considered reliable for statistical analysis.

**Variables:** **Obesity Rate (quantitative):** Percentage of adults classified as obese in each state, from the CDC.  
**Physical Activity (quantitative):** Percentage of adults meeting recommended physical activity guidelines, from the CDC.  
**Median Income (quantitative):** Median household income for each state (in U.S. dollars), from the U.S. Census Bureau.  
**College Degree (quantitative):** Percentage of adults with a bachelor's degree or higher, from the U.S. Census Bureau.  
**Fast Food per 100k (quantitative):** Number of fast food restaurants per 100,000 residents, from public health and food environment databases.  
**Fruit & Vegetable Consumption (quantitative):** Percentage of adults consuming recommended daily servings, from CDC nutrition data.  
**Region (categorical):** U.S. geographic region (South, Midwest, West, Northeast), based on U.S. Census classifications.

## Descriptive Statistics:

Obesity rates across states were approximately normally distributed with a slight right skew, indicating that a few states had higher-than-average obesity rates. Most states clustered around the mean, and no extreme outliers were observed, although a small number of states appeared on the higher end of the distribution.  $M = 35.26$ ,  $SD = 3.87$ ,  $Range = [25.1, 41.5]$

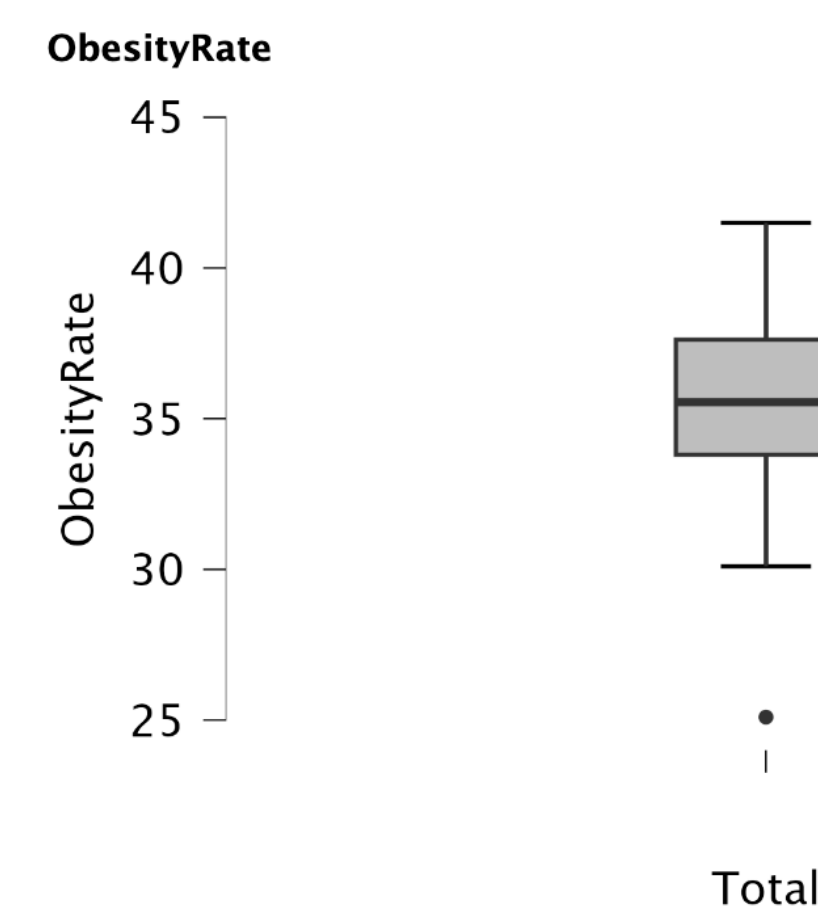


Fig 1. Box plot of Obesity Rate

Physical activity levels were fairly normally distributed with a slight left skew, suggesting that more states reported moderate to higher levels of physical activity, while fewer states reported very low activity levels. No major outliers were present.  $M = 53.60$ ,  $SD = 6.44$ ,  $Range = [44, 69]$

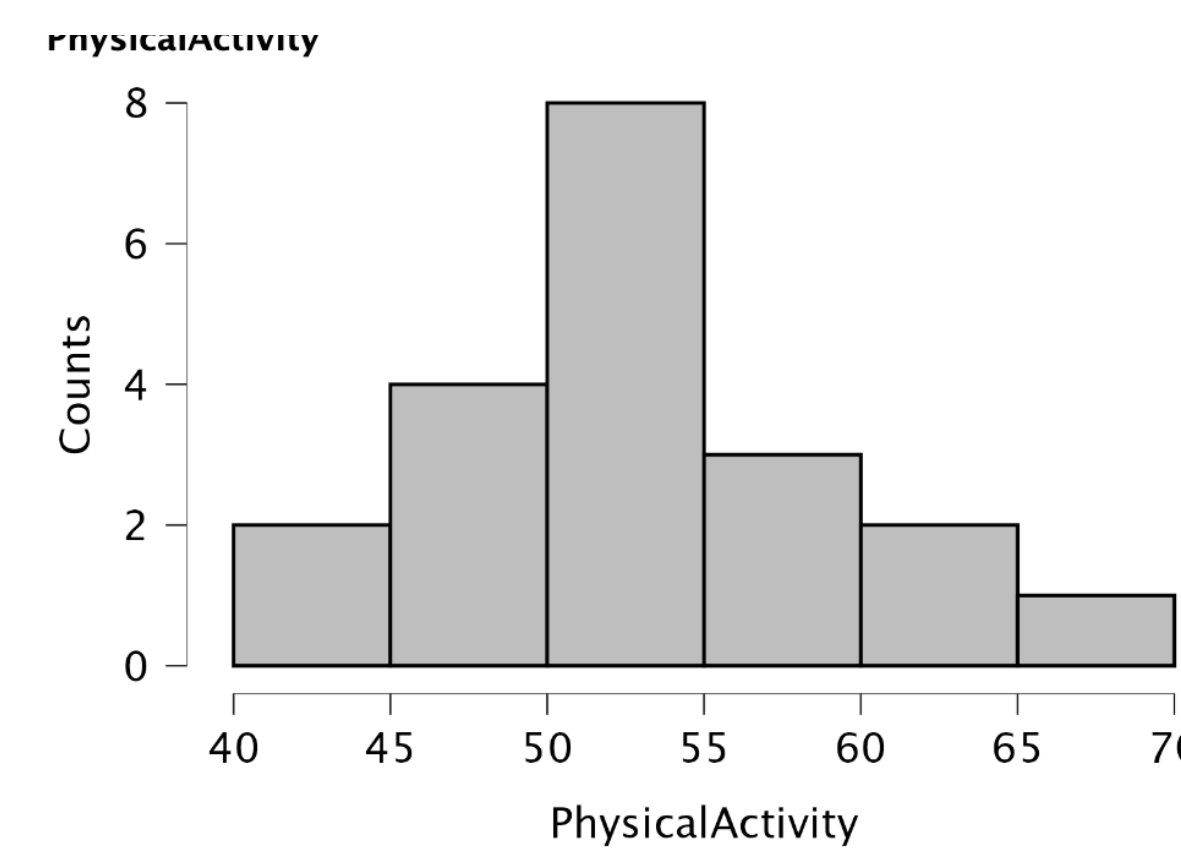


Fig 2. Histogram of Physical Activity

Median income values showed moderate variability across states and appeared slightly right-skewed, indicating that a few states have higher income levels compared to the majority. Most states cluster around the mean, and no extreme outliers were observed.  $M = 68,580$ ,  $SD = 11,030$ ,  $Range = [49,210, 84,950]$ .

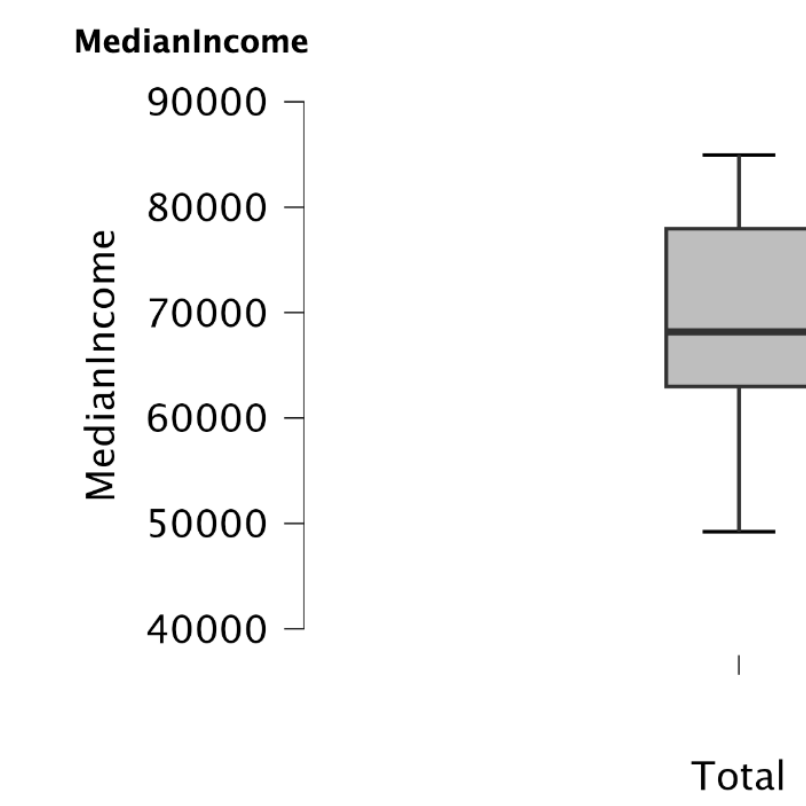


Fig 3. Box plot of Median Income

States were unevenly distributed across geographic regions, with the largest proportion located in the South, followed by the Midwest and West, and only one state in the Northeast. South:  $n = 9$ , 45.0%. Midwest:  $n = 6$ , 30.0%. West:  $n = 4$ , 20.0%. Northeast:  $n = 1$ , 5.0%

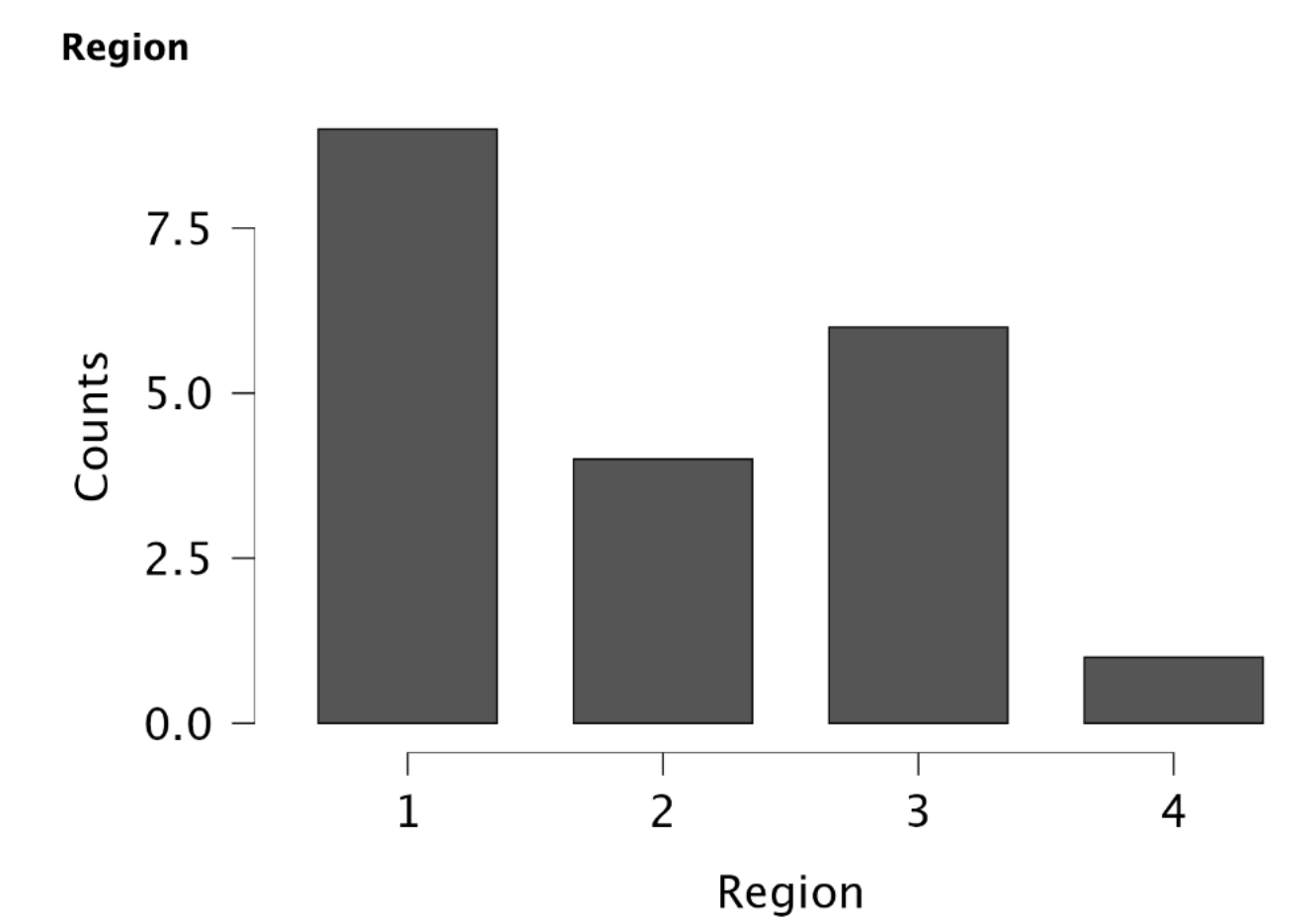


Fig 4. Bar graph of Regions

## Inferential Test 1: Linear Regression

A linear regression analysis was conducted to examine whether physical activity predicts obesity rates across U.S. states. This test was appropriate because both variables are continuous, and the goal was to determine the strength and direction of their relationship. It was expected that higher physical activity would be associated with lower obesity rates.

Results indicated that physical activity was a significant predictor of obesity rates, with a strong negative relationship. States with higher levels of physical activity tended to have lower obesity prevalence. The model explained a large proportion of the variance in obesity rates, indicating a strong overall fit.  $R^2 = .941$ ,  $F(1,18) = 288.10$ ,  $p < .001$ .

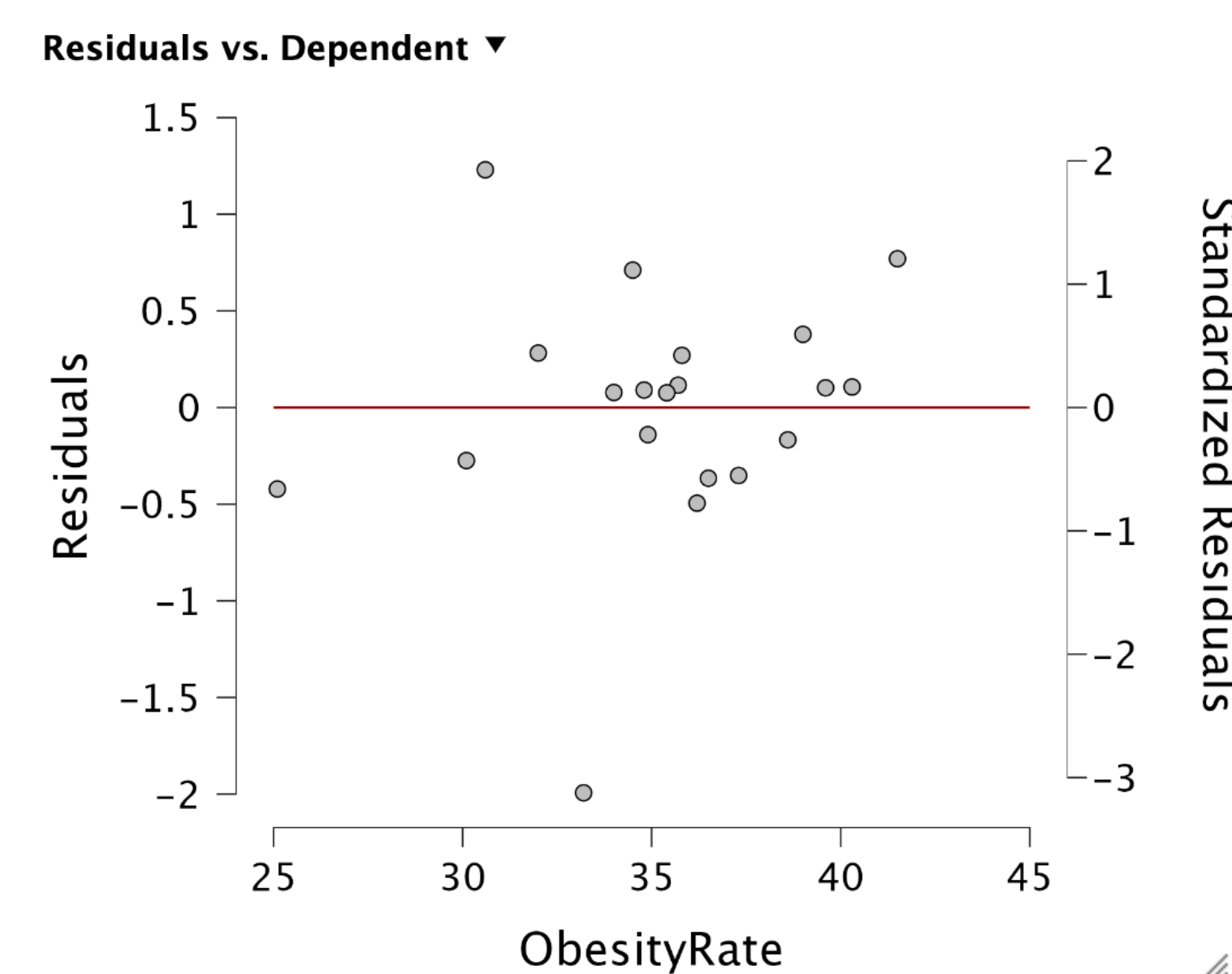


Fig 5. Scatter plot of obesity rate vs. physical activity

## Inferential Test 2: Independent Samples T-test

An independent samples t-test was conducted to compare obesity rates between states with high and low levels of physical activity. This test was appropriate because it compares the means of two independent groups. It was expected that states with higher physical activity would have lower obesity rates.

The results showed a significant difference between groups, with states in the high physical activity group having lower obesity rates than those in the low activity group. The effect size was very large, indicating a meaningful difference between the groups.  $t(18) = -4.69$ ,  $p < .001$ ,  $d = -2.10$ .

### Independent Samples T-Test

	t	df	p	Cohen's d	SE Cohen's d
ObesityRate	-4.685	18	< .001	-2.095	0.648

Note. Student's t-test.

Fig 6. Independent Samples T-test Results

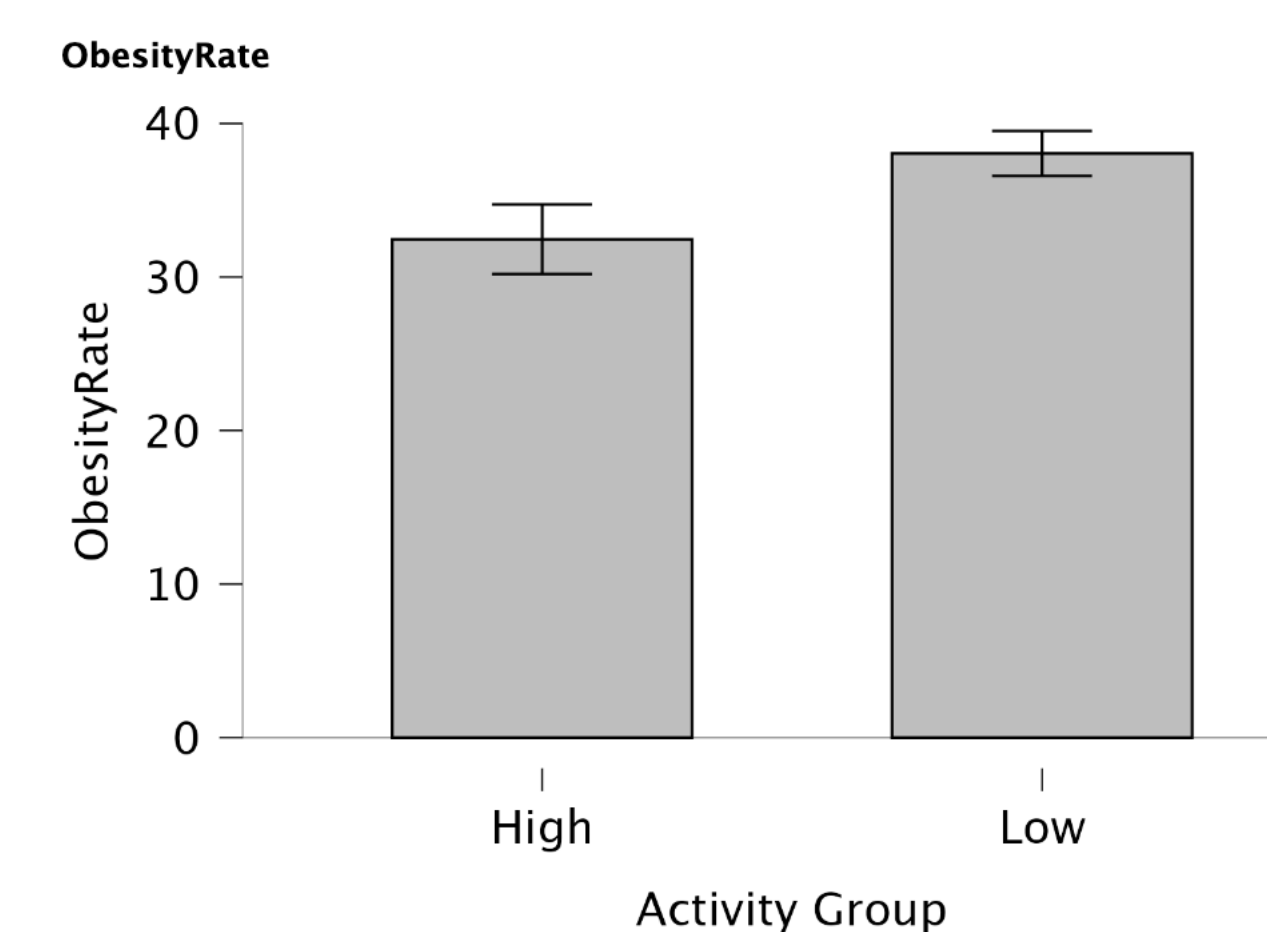


Fig 7. Mean obesity rates for high and low activity groups