**Race, Height, Weight, and Physical Activity: Comparing Variables**

**Jose Vazquez**

**Abstract**

This is a collection of arbitrary comparisons between two groupings of people defined principally by their skin's hue. Who is taller, blacks or whites? Who weighs more on average?

Which one engages in more physical activity every week? Additionally, I look at whether the weight has any correlation with household income. Is there a statistically significant difference in any of these measures between the races? Do people who bring in more revenue weigh less on average? Depending on how many black and white adults you've encountered, you might have guessed that the data would indicate a difference in heights between the groups, a difference in weights, or that there is no difference in average weekly physical activity. On the other hand, you may have also had no idea that whites were on average taller (mean difference= 0.98cm; 95% CI: (0.26, 1.70); P<0.05) or that blacks weigh more on average (mean difference= 5.42cm; 95% CI: (3.59, 7.25); P<0.05), or that average physical activity was independent of race (mean difference= 0.022; 95% CI (-0.12, 0.16); P> 0.05). My comparisons are prime illustrations that our intuition may not serve us well when encountering all statistical questions. Our life experience on such questions is often equivalent to a small and convenient sample prone to high variation and bias. However, it highlights that our reason may always come in handy, at most to unravel the ball of mixed yarn presented in some study and at least to recognize that almost anything can be compared and quantified, with little to no merit behind the comparison. And that arbitrary opinions can often be formed to fuel biases or prejudices under the guise of scientific analysis. In other words, anyone, including undergraduate students, lacking quantitative competency, severely lacking in coding ability, and with dilettante’s knowledge of statistics can present a visual comparison of a few variables.

**Introduction/ Objective**

Although Statistics is a great tool for understanding measurement error or variation as a natural byproduct of measurement, it is also the science of collecting information in large enough quantities for inferring something of a whole from a sample of that whole. It says little to nothing of how we might define a whole or category. You could imagine that rather than skin color, an apartheid society might separate its constituents by whether they fit the agreed-upon definition of tall by an agreed-upon age. In many instances, the definitions may have substantial reasoning behind them, but again, it would be up to society to decide if these reasons carry any clout or merit. There are some truths science can point us towards, and there are others for which it can offer no answer. Like reason, the science of statistics is a tool. We often use tools to pursue an inclination, presupposition, attempt to corroborate a bias, and the like. Therefore, we should be keenly aware of our beliefs to best safeguard against the misapplication or unintended application of said tools. I use statistics and a few borrowed lines of code as tools to compare blacks with whites in height, weight, and weekly physical activity level, and I see if there is a relationship between weight and household income. The purpose of this project is to present visualizations of these variables. And to pose the question of whether there are any meaningful conclusions, decisions, or opinions derived from the statistical analysis. I suspect that, however arbitrary these comparisons may seem, someone somewhere might use this to corroborates another belief. On the other hand, it's not too much for me to assume that most people will not find that this information has any real consequences to their knowledge base or opinions.

**Materials and Methods**

**Statistical Analysis**

The analysis was done on a Python simulator, Codio.com.

Additional packages used were: Numpy, Pandas, and Matplotlib.

<https://github.com/>

I look at adult height data of black and white heights from the National Health and Nutrition Examination Survey. I defined adults as being 21 years or older.

**Dataset and Codes**

My project's data comes from the National Health and Nutrition Examination Survey (NHANES) 2011 – 2018 dataset.

My codes primarily come from lecture notes and homework assignments for my statistics course.

All credit for codes goes to Diaki Tagami.

**Analysis and Results**



First, I looked at the relationship between **height** and **race** variables. I found that the mean height of adult **blacks**(21 years and older) is 169.0cm, with 95% confidence that adult blacks' average population height is between 168.3cm and 169.7cm. The mean height of adult **whites** (21 years and older) is 170.0cm, with a 95% confidence that the population mean between 169.7cm and 170.3cm. The average difference (white average - black average) is about 1.0cm, with a 95% confidence that the population difference lies between 0.26cm and 1.71cm. Along with the mean confidence intervals, I compared the mean difference of the two groups using a t-test. A t-test gives us a standardized score that tells us how far we expect our sample mean difference to deviate from a null value of population mean difference, set at zero, to indicate no difference. The test statistic resulted in 2.66 and a p-value of 0.00785. Because our test statistic is more than two standard errors away, by 0.66, from our null value of zero, indicating no difference. That value is considered too extreme to be attributed by chance alone, which in turn gives us a p-value of less than 0.05 and allows us to claim a statistically significant difference between the heights of blacks and whites. I can then say I have found a statistical significance that indicates that individuals in the white group are, on average taller than those in the black group.



Next, I looked at the relationship between w**eight** and **race** variables. I found that the mean weight of adult **blacks**(21 years and older) is 88.32kg, with 95% confidence that adult blacks' average population weight is between 86.58 kg and 90.05 kg. The mean weight of adult **whites** (21 years and older) is 82.90 kg, with a 95% confidence that the population mean between 82.31 kg and 83.49 kg. The average weight difference is about 5.42 kg, with a 95% confidence that the population difference lies between 3.59 kg and 7.25 kg. The test statistic resulted in 5.77 and a p-value of 0.000000010. Since the test statistic is more than two standard errors away from our null value of zero, the value is considered too extreme to be attributed by chance alone. This, in turn, gives us a p-value of less than 0.05 and allows us to claim a statistically significant difference between the weights of blacks and whites. I can then say I have found a statistical significance that indicates that individuals in the white group are, on average taller than those in the black group.



Thirdly, I looked at the relationship between weekly physical activity (WPA) and **race** variables. I found that the mean WPA of adult **blacks**is 3.68 days, with 95% confidence that the average population WPA of adult blacks is between 3.55 days and 3.81 days. The mean WPA of adult **whites** is 3.70 days, with a 95% confidence that the population mean between 3.65 days and 3.75 days. The average WPA difference is about 0.02 days, with a 95% confidence that the population difference lies between -0.12 days and 0.16 days. The test statistic resulted in -0.21 and a p-value of 0.83. That value falls within our minus two standard error borderline, which gives us a p-value of more than 0.05 and does not allow us to reject our null value or claim a statistically significant difference between the WPA of blacks and whites.



Lastly, I looked at whether there was any correlation between household income and weight. The Prob (F-statistic) is 0.74. The correlation coefficient is near zero 0.0040 and a p-value equal to 0.74, denying us the ability to claim statistical significance. No correlation was found.

**Conclusion**

My statistical analysis provided evidence of racial difference in height and weight, none in physical activity, and no correlation between income and weight. These arbitrary comparisons of variables are to illustrate that visual quantifications can be performed on any number of variables.

**Credits**

https://www.rdocumentation.org/packages/NHANES/versions/2.1.0/topics/NHANES