

Body Mass Index Survey Analysis

Introduction

The Body Mass Index is an important factor as it is widely regarded to the chances of having a longer and healthier life if you have an ideal index. Basically, is a way to know if your weight is proportional to your height and can help you to determine any health risks you may face if it's outside the healthy range.

Objective

In this project we will analyze the NHANES survey and use a “survey-weighted” regression to test any relationships between measurements. We will gather the attention around the Body Mass Index (BMI) and its relationship with physical activity.

Data

A good source to analyze this factor is the National Health and Nutrition Examination Survey (NHANES). It is a complex survey designed to measure the population's health and nutritional status in the United States. The data includes measurements related to overall health, diet, physical activity, mental health, socioeconomic factors, etc, of over 20,000 individuals including adults and children.

```
library(NHANES)
library(dplyr)
library(ggplot2)
library(survey)
library(broom)
library(quantreg)

data("NHANESraw")
df <- NHANESraw
glimpse(df)
```

```
## Rows: 20,293
## Columns: 78
## $ ID          <int> 51624, 51625, 51626, 51627, 51628, 51629, 51630, 5...
## $ SurveyYr    <fct> 2009_10, 2009_10, 2009_10, 2009_10, 2009_10, 2009_...
## $ Gender      <fct> male, male, male, male, female, male, female, fema...
## $ Age         <int> 34, 4, 16, 10, 60, 26, 49, 1, 10, 80, 10, 80, 4, 3...
## $ AgeMonths   <int> 409, 49, 202, 131, 722, 313, 596, 12, 124, NA, 121...
## $ Race1       <fct> White, Other, Black, Black, Black, Mexican, White,...
## $ Race3       <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA...
## $ Education   <fct> High School, NA, NA, NA, High School, 9 - 11th Gra...
## $ MaritalStatus <fct> Married, NA, NA, NA, Widowed, Married, LivePartner...
## $ HHIncome    <fct> 25000-34999, 20000-24999, 45000-54999, 20000-24999...
## $ HHIncomeMid <int> 30000, 22500, 50000, 22500, 12500, 30000, 40000, 4...
## $ Poverty     <dbl> 1.36, 1.07, 2.27, 0.81, 0.69, 1.01, 1.91, 1.36, 2...
## $ HomeRooms   <int> 6, 9, 5, 6, 6, 4, 5, 5, 7, 4, 5, 5, 7, NA, 6, 6, 5...
## $ HomeOwn     <fct> Own, Own, Own, Rent, Rent, Rent, Rent, Rent, Own, ...
## $ Work        <fct> NotWorking, NA, NotWorking, NA, NotWorking, Workin...
## $ Weight      <dbl> 87.4, 17.0, 72.3, 39.8, 116.8, 97.6, 86.7, 9.4, 26...
## $ Length      <dbl> NA, NA, NA, NA, NA, NA, NA, 75.7, NA, NA, NA, NA, ...
## $ HeadCirc    <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA...
```

```

## $ Height <dbl> 164.7, 105.4, 181.3, 147.8, 166.0, 173.0, 168.4, N...
## $ BMI <dbl> 32.22, 15.30, 22.00, 18.22, 42.39, 32.61, 30.57, N...
## $ BMICatUnder20yrs <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA...
## $ BMI_WHO <fct> 30.0_plus, 12.0_18.5, 18.5_to_24.9, 12.0_18.5, 30....
## $ Pulse <int> 70, NA, 68, 68, 72, 72, 86, NA, 70, 88, 84, 54, NA...
## $ BPSysAve <int> 113, NA, 109, 93, 150, 104, 112, NA, 108, 139, 94,...
## $ BPDiaAve <int> 85, NA, 59, 41, 68, 49, 75, NA, 53, 43, 45, 60, NA...
## $ BPSys1 <int> 114, NA, 112, 92, 154, 102, 118, NA, 106, 142, 94,...
## $ BPDia1 <int> 88, NA, 62, 36, 70, 50, 82, NA, 60, 62, 38, 62, NA...
## $ BPSys2 <int> 114, NA, 114, 94, 150, 104, 108, NA, 106, 140, 92,...
## $ BPDia2 <int> 88, NA, 60, 44, 68, 48, 74, NA, 50, 46, 40, 62, NA...
## $ BPSys3 <int> 112, NA, 104, 92, 150, 104, 116, NA, 110, 138, 96,...
## $ BPDia3 <int> 82, NA, 58, 38, 68, 50, 76, NA, 56, 40, 50, 58, NA...
## $ Testosterone <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA...
## $ DirectChol <dbl> 1.29, NA, 1.55, 1.89, 1.16, 1.16, 1.16, NA, 1.58, ...
## $ TotChol <dbl> 3.49, NA, 4.97, 4.16, 5.22, 4.14, 6.70, NA, 4.14, ...
## $ UrineVol1 <int> 352, NA, 281, 139, 30, 202, 77, NA, 39, 128, 109, ...
## $ UrineFlow1 <dbl> NA, NA, 0.415, 1.078, 0.476, 0.563, 0.094, NA, 0.3...
## $ UrineVol2 <int> NA, NA, NA, NA, 246, NA, NA, NA, NA, NA, NA, NA, N...
## $ UrineFlow2 <dbl> NA, NA, NA, NA, 2.51, NA, NA, NA, NA, NA, NA, NA, ...
## $ Diabetes <fct> No, No, No, No, Yes, No, No, No, No, No, No, Yes, ...
## $ DiabetesAge <int> NA, NA, NA, NA, 56, NA, NA, NA, NA, NA, NA, 70, NA...
## $ HealthGen <fct> Good, NA, Vgood, NA, Fair, Good, Good, NA, NA, Exc...
## $ DaysPhysHlthBad <int> 0, NA, 2, NA, 20, 2, 0, NA, NA, 0, NA, 0, NA, NA, ...
## $ DaysMentHlthBad <int> 15, NA, 0, NA, 25, 14, 10, NA, NA, 0, NA, 0, NA, N...
## $ LittleInterest <fct> Most, NA, NA, NA, Most, None, Several, NA, NA, Non...
## $ Depressed <fct> Several, NA, NA, NA, Most, Most, Several, NA, NA, ...
## $ nPregnancies <int> NA, NA, NA, NA, 1, NA, 2, NA, NA, NA, NA, NA, NA, ...
## $ nBabies <int> NA, NA, NA, NA, 1, NA, 2, NA, NA, NA, NA, NA, NA, ...
## $ Age1stBaby <int> NA, NA, NA, NA, NA, NA, 27, NA, NA, NA, NA, NA, NA...
## $ SleepHrsNight <int> 4, NA, 8, NA, 4, 4, 8, NA, NA, 6, NA, 9, NA, 7, NA...
## $ SleepTrouble <fct> Yes, NA, No, NA, No, No, Yes, NA, NA, No, NA, No, ...
## $ PhysActive <fct> No, NA, Yes, NA, No, Yes, No, NA, NA, Yes, NA, No,...
## $ PhysActiveDays <int> NA, NA, 5, NA, NA, 2, NA, NA, NA, 4, NA, NA, NA, N...
## $ TVHrsDay <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA...
## $ CompHrsDay <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA...
## $ TVHrsDayChild <int> NA, 4, NA, 1, NA, NA, NA, NA, 1, NA, 3, NA, 2, NA,...
## $ CompHrsDayChild <int> NA, 1, NA, 1, NA, NA, NA, NA, 0, NA, 0, NA, 1, NA,...
## $ Alcohol12PlusYr <fct> Yes, NA, NA, NA, No, Yes, Yes, NA, NA, Yes, NA, No...
## $ AlcoholDay <int> NA, NA, NA, NA, NA, 19, 2, NA, NA, 1, NA, NA, NA, ...
## $ AlcoholYear <int> 0, NA, NA, NA, 0, 48, 20, NA, NA, 52, NA, 0, NA, N...
## $ SmokeNow <fct> No, NA, NA, NA, Yes, No, Yes, NA, NA, No, NA, No, ...
## $ Smoke100 <fct> Yes, NA, NA, NA, Yes, Yes, Yes, NA, NA, Yes, NA, Y...
## $ SmokeAge <int> 18, NA, NA, NA, 16, 15, 38, NA, NA, 16, NA, 21, NA...
## $ Marijuana <fct> Yes, NA, NA, NA, NA, Yes, Yes, NA, NA, NA, NA, NA,...
## $ AgeFirstMarij <int> 17, NA, NA, NA, NA, 10, 18, NA, NA, NA, NA, NA, NA...
## $ RegularMarij <fct> No, NA, NA, NA, NA, Yes, No, NA, NA, NA, NA, NA, NA, N...
## $ AgeRegMarij <int> NA, NA, NA, NA, NA, 12, NA, NA, NA, NA, NA, NA, NA...
## $ HardDrugs <fct> Yes, NA, NA, NA, No, Yes, Yes, NA, NA, NA, NA, NA,...
## $ SexEver <fct> Yes, NA, NA, NA, Yes, Yes, Yes, NA, NA, NA, NA, NA...
## $ SexAge <int> 16, NA, NA, NA, 15, 9, 12, NA, NA, NA, NA, NA, NA,...
## $ SexNumPartnLife <int> 8, NA, NA, NA, 4, 10, 10, NA, NA, NA, NA, NA, NA, ...
## $ SexNumPartYear <int> 1, NA, NA, NA, NA, 1, 1, NA, NA, NA, NA, NA, NA, N...
## $ SameSex <fct> No, NA, NA, NA, No, No, Yes, NA, NA, NA, NA, NA, N...

```

```
## $ SexOrientation    <fct> Heterosexual, NA, NA, NA, NA, Heterosexual, Hetero...
## $ WTINT2YR          <dbl> 80100.544, 53901.104, 13953.078, 11664.899, 20090....
## $ WTMEC2YR          <dbl> 81528.772, 56995.035, 14509.279, 12041.635, 21000....
## $ SDMVPSU           <int> 1, 2, 1, 2, 2, 1, 2, 2, 2, 1, 1, 1, 2, 2, 1, 1, 1,...
## $ SDMVSTRA           <int> 83, 79, 84, 86, 75, 88, 85, 86, 88, 77, 86, 79, 84...
## $ PregnantNow        <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA...
```

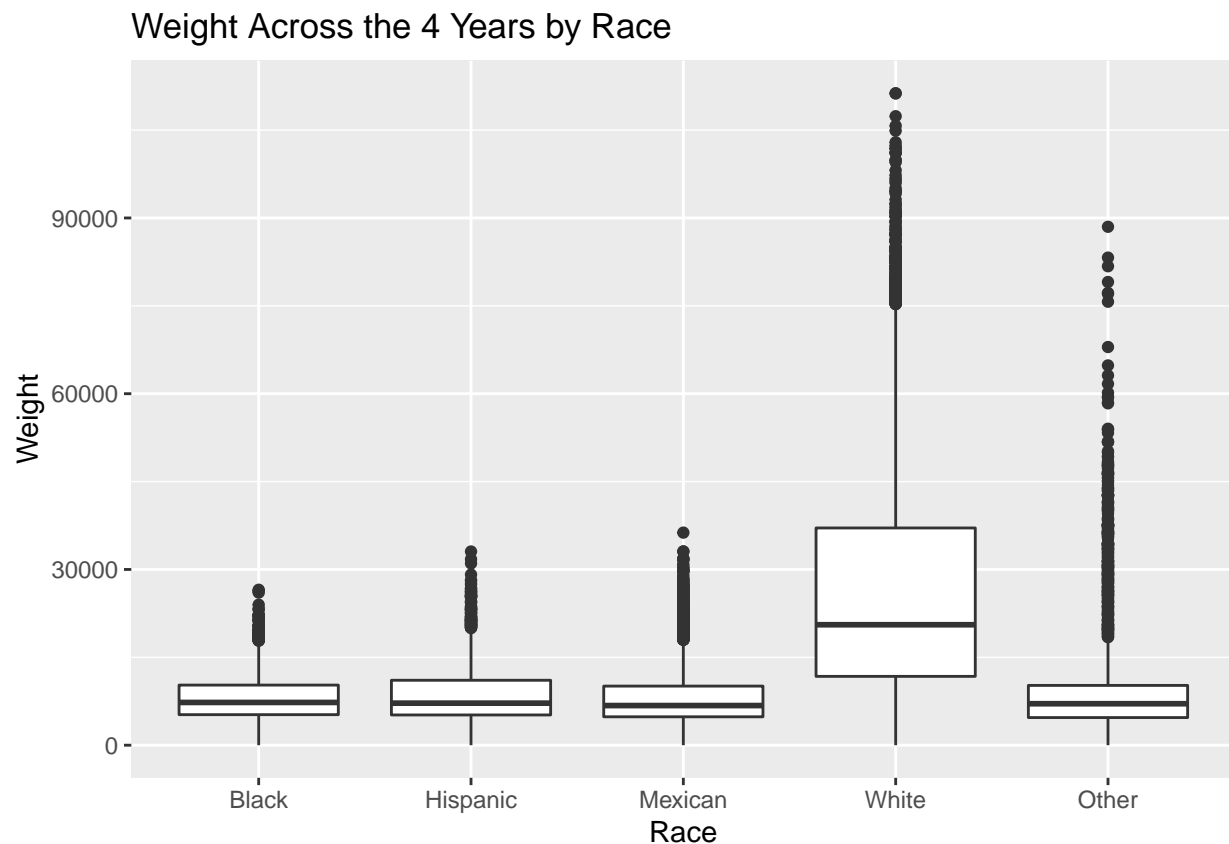
By having a glimpse to the data we can realize that it contains 4 years of historical data (2009-2012) and that contains data of different regions that include different races such as Mexicans, Afroamericans, Hispanic, etc.

Exploratory Analysis

Lets have a look to the weights by race based on the 4 years of data. Currently the weights sum to 2 times the US population number, so we need to divide the 2-year weight in half so that in total, the sum of the weights is equal to the US population.

```
df <- df %>% mutate(WTMEC4YR = WTMEC2YR/2)

ggplot(df, aes(x = Race1, y = WTMEC4YR)) +
  geom_boxplot() +
  labs(title='Weight Across the 4 Years by Race') +
  xlab('Race') +
  ylab('Weight')
```



As you can see there is a higher tendency of the weight across the 4 years for the white race compared to the others.

Using the survey library we can specify the survey design for analyses purposes. We need to specify the design so the sampling weights and design are used properly in the statistical models. We will focus into 2 particular variables for design effects of stratification and clustering.

```
df_design <- svydesign(
  data = df,
  strata = ~SDMVSTRA,
  id = ~SDMVPSU,
  nest = TRUE,
  weights = ~WTMEC4YR)

summary(df_design)
```

```
## Stratified 1 - level Cluster Sampling design (with replacement)
## With (62) clusters.
## svydesign(data = df, strata = ~SDMVSTRA, id = ~SDMVPSU, nest = TRUE,
##           weights = ~WTMEC4YR)
## Probabilities:
##      Min.   1st Qu.   Median     Mean   3rd Qu.     Max.
## 8.986e-06 5.664e-05 1.054e-04      Inf 1.721e-04      Inf
## Stratum Sizes:
##           75  76  77  78  79  80  81  82  83  84  85  86  87  88  89  90  91
## obs           803 785 823 829 696 751 696 724 713 683 592 946 598 647 251 862 998
## design.PSU     2   2   2   2   2   2   2   2   2   2   2   3   2   2   2   3   3
## actual.PSU     2   2   2   2   2   2   2   2   2   2   2   3   2   2   2   3   3
##           92  93  94  95  96  97  98  99 100 101 102 103
## obs           875 602 688 722 676 608 708 682 700 715 624 296
## design.PSU     3   2   2   2   2   2   2   2   2   2   2   2
## actual.PSU     3   2   2   2   2   2   2   2   2   2   2   2
## Data variables:
## [1] "ID"                "SurveyYr"          "Gender"            "Age"
## [5] "AgeMonths"         "Race1"             "Race3"             "Education"
## [9] "MaritalStatus"     "HHIncome"          "HHIncomeMid"       "Poverty"
## [13] "HomeRooms"         "HomeOwn"           "Work"              "Weight"
## [17] "Length"            "HeadCirc"          "Height"            "BMI"
## [21] "BMICatUnder20yrs"  "BMI_WHO"           "Pulse"             "BPSysAve"
## [25] "BPDiaAve"          "BPSys1"            "BPDia1"            "BPSys2"
## [29] "BPDia2"            "BPSys3"            "BPDia3"            "Testosterone"
## [33] "DirectChol"        "TotChol"           "UrineVol1"         "UrineFlow1"
## [37] "UrineVol2"         "UrineFlow2"        "Diabetes"          "DiabetesAge"
## [41] "HealthGen"         "DaysPhysHlthBad"   "DaysMentHlthBad"   "LittleInterest"
## [45] "Depressed"         "nPregnancies"      "nBabies"           "Age1stBaby"
## [49] "SleepHrsNight"     "SleepTrouble"      "PhysActive"        "PhysActiveDays"
## [53] "TVHrsDay"          "CompHrsDay"        "TVHrsDayChild"     "CompHrsDayChild"
## [57] "Alcohol12PlusYr"   "AlcoholDay"        "AlcoholYear"       "SmokeNow"
## [61] "Smoke100"          "SmokeAge"          "Marijuana"         "AgeFirstMarij"
## [65] "RegularMarij"      "AgeRegMarij"       "HardDrugs"         "SexEver"
## [69] "SexAge"            "SexNumPartnLife"   "SexNumPartYear"    "SameSex"
## [73] "SexOrientation"    "WTINT2YR"          "WTMEC2YR"          "SDMVPSU"
## [77] "SDMVSTRA"          "PregnantNow"       "WTMEC4YR"
```

Body mass index categories are distinct for children and adults, so let's subset the data to only analyze adults of at least 20 years old.

```
df_adult <- subset(df_design, Age >= 20)
summary(df_adult)
```

```
## Stratified 1 - level Cluster Sampling design (with replacement)
## With (62) clusters.
## subset(df_design, Age >= 20)
## Probabilities:
##      Min.   1st Qu.   Median     Mean   3rd Qu.     Max.
## 8.986e-06 4.303e-05 8.107e-05      Inf 1.240e-04      Inf
## Stratum Sizes:
##      75  76  77  78  79  80  81  82  83  84  85  86  87  88  89  90  91
## obs      471 490 526 500 410 464 447 400 411 395 357 512 327 355 153 509 560
## design.PSU  2  2  2  2  2  2  2  2  2  2  2  3  2  2  2  3  3
## actual.PSU  2  2  2  2  2  2  2  2  2  2  2  3  2  2  2  3  3
##      92  93  94  95  96  97  98  99 100 101 102 103
## obs      483 376 368 454 362 315 414 409 377 460 308 165
## design.PSU  3  2  2  2  2  2  2  2  2  2  2  2
## actual.PSU  3  2  2  2  2  2  2  2  2  2  2  2
## Data variables:
## [1] "ID"                "SurveyYr"            "Gender"              "Age"
## [5] "AgeMonths"          "Race1"               "Race3"              "Education"
## [9] "MaritalStatus"      "HHIncome"            "HHIncomeMid"         "Poverty"
## [13] "HomeRooms"          "HomeOwn"             "Work"               "Weight"
## [17] "Length"             "HeadCirc"            "Height"              "BMI"
## [21] "BMICatUnder20yrs"   "BMI_WHO"             "Pulse"               "BPSysAve"
## [25] "BPDiaAve"           "BPSys1"              "BPDia1"              "BPSys2"
## [29] "BPDia2"             "BPSys3"              "BPDia3"              "Testosterone"
## [33] "DirectChol"         "TotChol"             "UrineVol1"           "UrineFlow1"
## [37] "UrineVol2"          "UrineFlow2"          "Diabetes"            "DiabetesAge"
## [41] "HealthGen"          "DaysPhysHlthBad"     "DaysMentHlthBad"     "LittleInterest"
## [45] "Depressed"          "nPregnancies"        "nBabies"             "Age1stBaby"
## [49] "SleepHrsNight"      "SleepTrouble"        "PhysActive"          "PhysActiveDays"
## [53] "TVHrsDay"           "CompHrsDay"          "TVHrsDayChild"       "CompHrsDayChild"
## [57] "Alcohol112PlusYr"   "AlcoholDay"          "AlcoholYear"         "SmokeNow"
## [61] "Smoke100"           "SmokeAge"            "Marijuana"           "AgeFirstMarij"
## [65] "RegularMarij"       "AgeRegMarij"         "HardDrugs"           "SexEver"
## [69] "SexAge"             "SexNumPartnLife"     "SexNumPartYear"      "SameSex"
## [73] "SexOrientation"     "WTINT2YR"            "WTMEC2YR"            "SDMVPSU"
## [77] "SDMVSTRA"           "PregnantNow"         "WTMEC4YR"
```

```
print(nrow(df_design))
```

```
## [1] 20293
```

```
print(nrow(df_adult))
```

```
## [1] 11778
```

The process above it's important since we are using the sampling methods to estimate the true measurements distributions within the total population. In this case we want to estimate the average body mass index in the us adult population and get a visualization of the distribution.

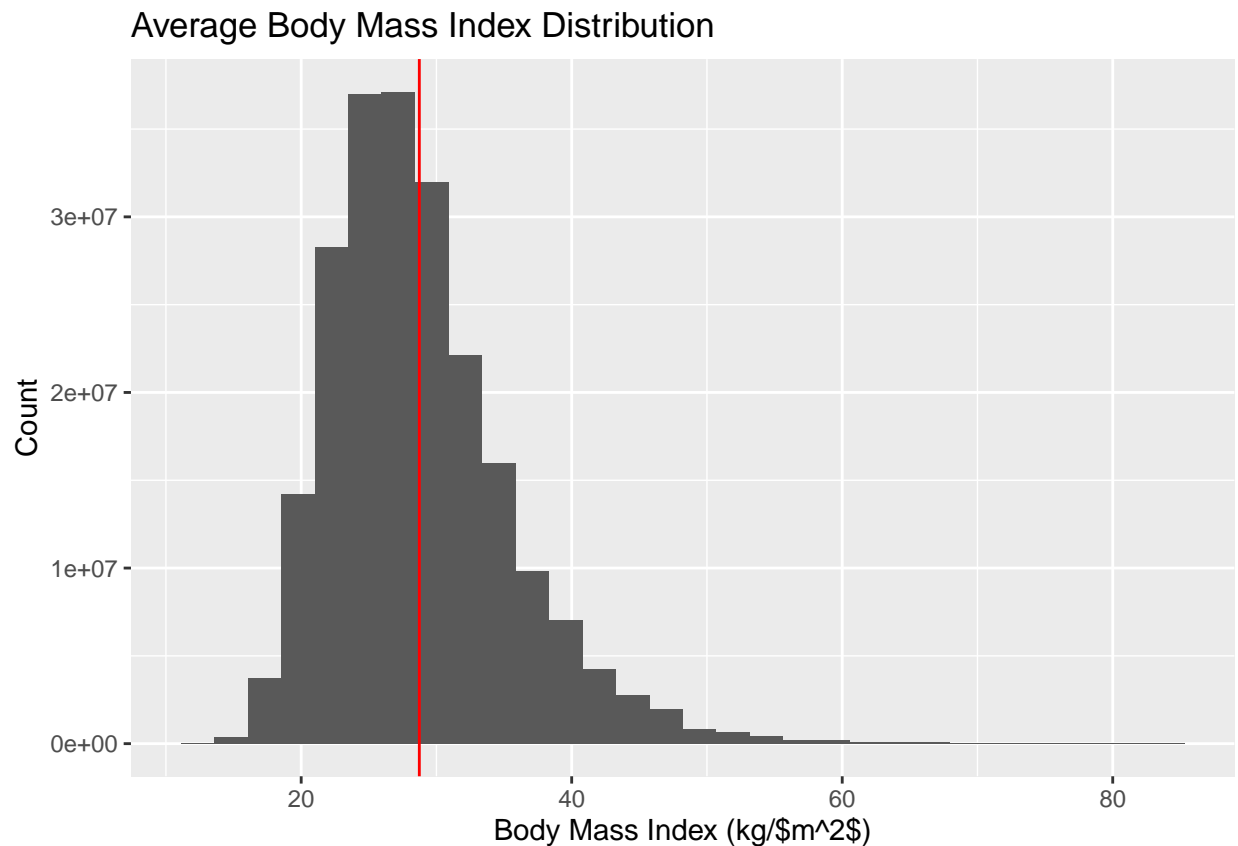
```
bmi_avgr <- df %>%
  filter(Age >= 20) %>%
  summarize(mean(BMI, na.rm=TRUE))
bmi_avgr
```

```
## # A tibble: 1 x 1
##   'mean(BMI, na.rm = TRUE)'
##                               <dbl>
## 1                               29.0
```

```
bmi_mean <- svymean(~BMI, design = df_adult, na.rm = TRUE)
bmi_mean
```

```
##      mean      SE
## BMI 28.734 0.1235
```

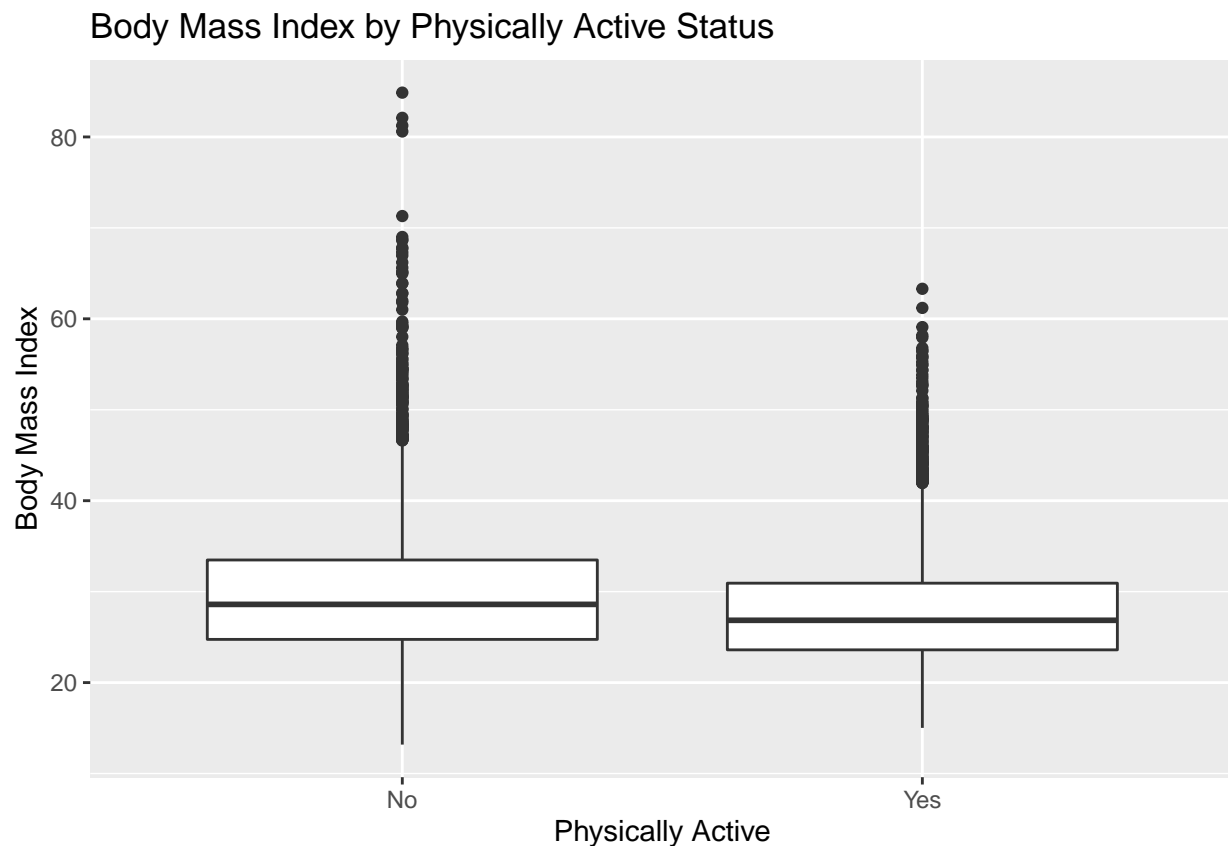
```
df %>%
  filter(Age >= 20) %>%
  ggplot(mapping = aes(x = BMI, weight = WTMEC4YR)) +
  geom_histogram() +
  geom_vline(xintercept = coef(bmi_mean), color="red") + labs(title='Average Body Mass Index Distribution')
```



The index distribution seems satisfactory since most people is under 40 kg/m² and even showing a slight skewness due to having a much higher index.

Now lets inquire if the distribution of body mass index differs between individual who are physically active against those who are not and compared them visually. Lets also make a t-test comparing the average index between physically active people.

```
df %>%
  filter(Age>=20) %>%
  ggplot(mapping = aes(x = PhysActive, y = BMI, weight = WTMEC4YR)) +
  geom_boxplot() + labs(title='Body Mass Index by Physically Active Status', x='Physically Active', y=
```



```
survey_ttest <- svytest(BMI~PhysActive, design = df_adult)
print(tidy(survey_ttest))
```

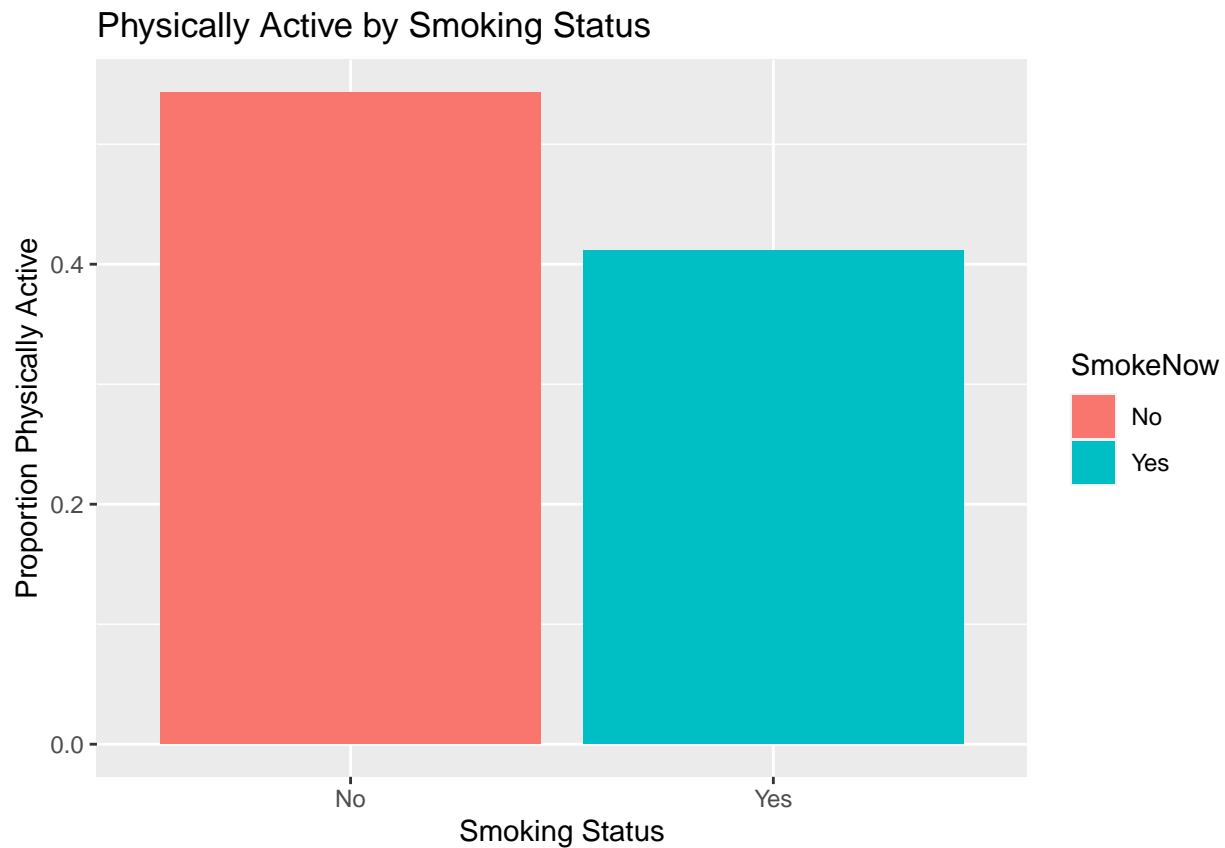
```
## # A tibble: 1 x 8
##   estimate statistic p.value parameter conf.low conf.high method alternative
##   <dbl>      <dbl>   <dbl>      <dbl>    <dbl>    <dbl> <chr>      <chr>
## 1    -1.85     -9.72 4.56e-11        32    -2.22    -1.47 Design-b~ two.sided
```

By the plot it may seem that people who are physically active have lower body mass indexes. That is partially true. It is actually more complex than a simple relation ship between this 2 factors, in fact there are more other factors that contribute to this relationship such as demographics, lifestyles and other particular ones.

Lets take a step ahead an investigate if smoking has a relationship with the body mass index.

```
phys_smoking <- svyby(~PhysActive, by = ~SmokeNow,
  FUN = svymean,
  design = df_adult,
  keep.names = FALSE)

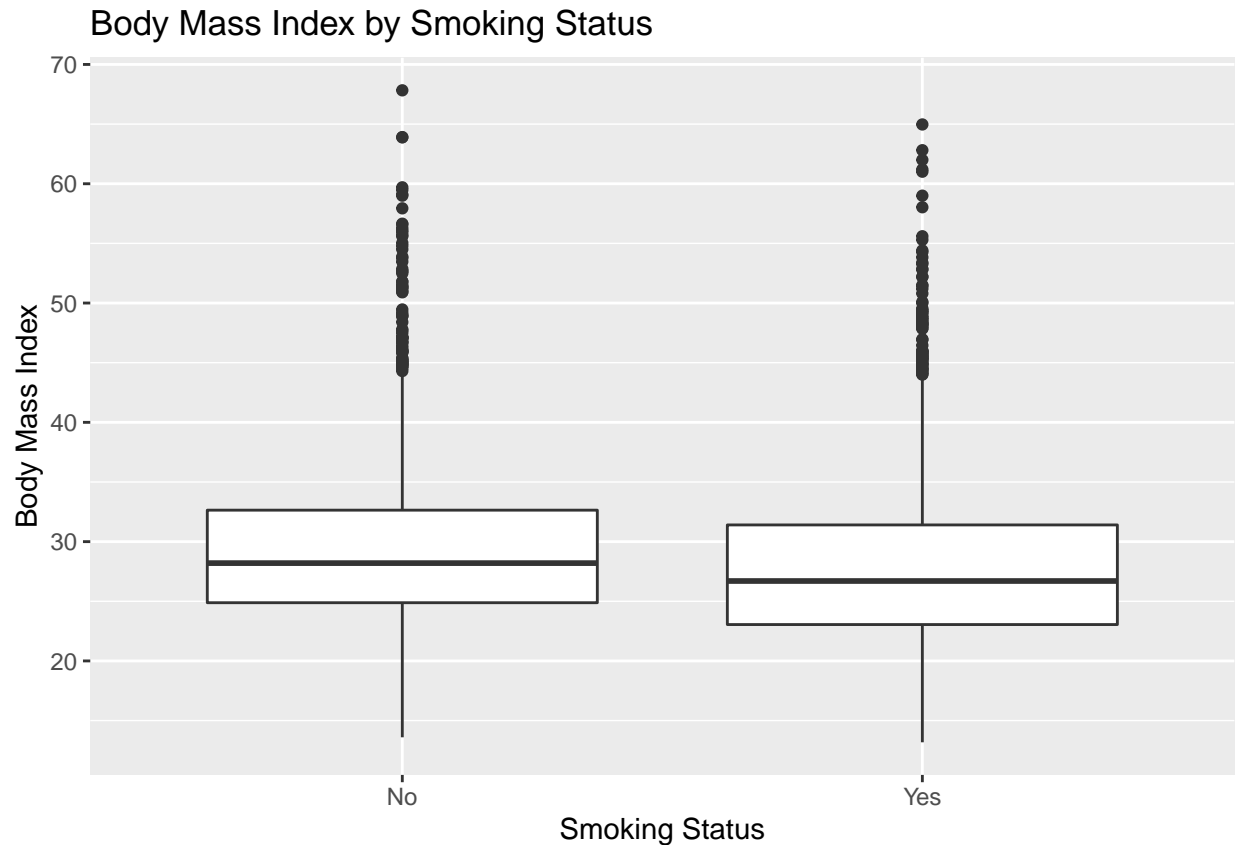
ggplot(data = phys_smoking,
  aes(y = PhysActiveYes, x = SmokeNow, fill = SmokeNow)) +
  geom_col() + labs(title='Physically Active by Smoking Status', x='Smoking Status', y="Proportion Ph
```



```
index_smoke <- svyby(~BMI, by = ~SmokeNow,
  FUN = svymean,
  design = df_adult,
  na.rm = TRUE)
index_smoke
```

```
##      SmokeNow      BMI      se
## No          No 29.25734 0.1915138
## Yes         Yes 27.74873 0.1652377
```

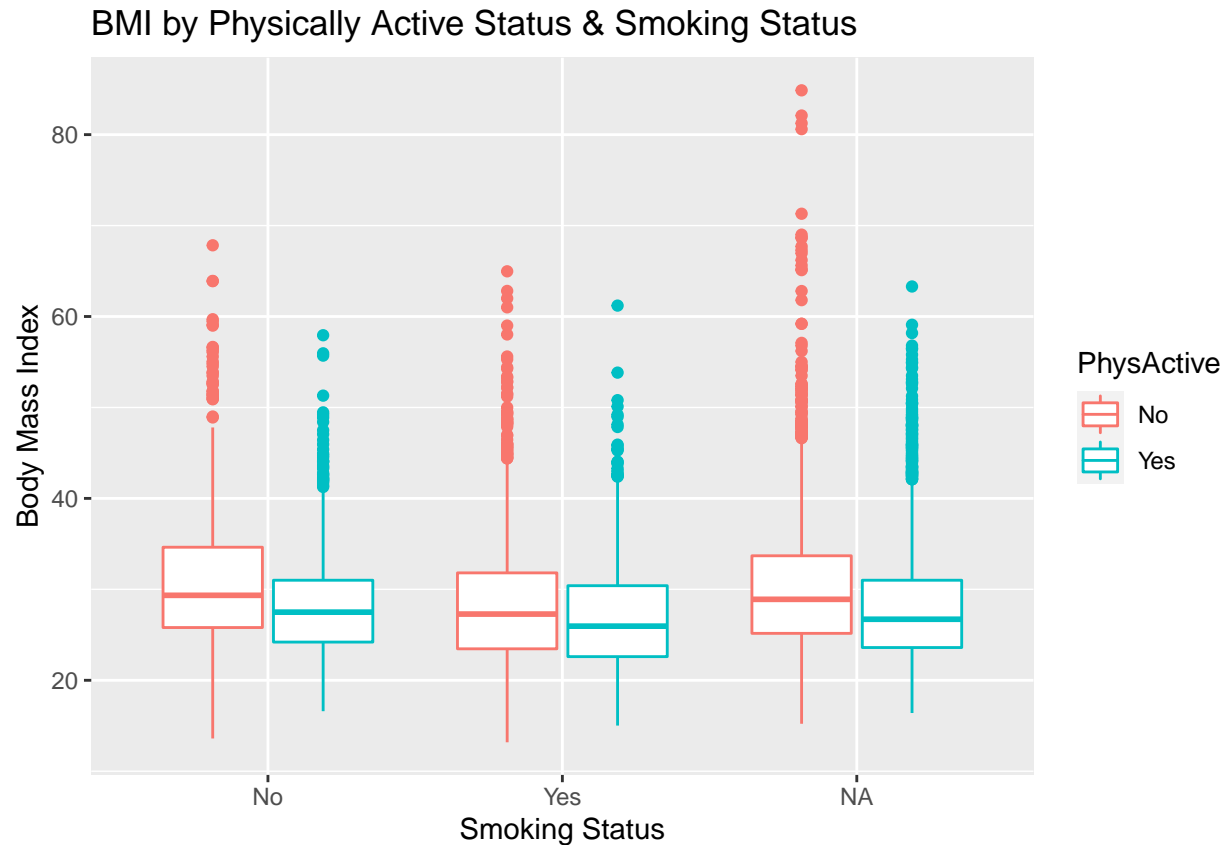
```
df %>%
  filter(Age>=20, !is.na(SmokeNow)) %>%
  ggplot(mapping = aes(x = SmokeNow, y = BMI, weight = WTMEC4YR)) +
  geom_boxplot() + labs(title='Body Mass Index by Smoking Status', x='Smoking Status', y="Body Mass I
```

Looking at these 2 charts we can say that people who smoke are less likely to be physically active and have a lower BMI on average. Additionally, people who are physically active have a lower body mass index on average.

To get a better insight of the situation, we can compare body mass index by physical activity stratified by smoking status in the chart below.

```
df %>%
  filter(Age>=20) %>%
  ggplot(mapping = aes(x = SmokeNow,
                        y = BMI,
                        weight = WTMEC4YR,
                        color = PhysActive)) +
  geom_boxplot() + labs(title='BMI by Physically Active Status & Smoking Status', x='Smoking Status',
```



Like previously seen, individual who are physically active tend to have a lower body mass index no matter their smoking status, same case for those who omitted answering the question in the survey. Interesting how smokers also have lower indexes and also the difference of index of the physically active to non active is slightly smaller in smokers than in no smokers.

Regression Model

As previously done, we are adding the smoking status as some toher possible predictors of bosy mass index by using a linear regression model with multiple independent variables. In this case we are going to use a weighted method since we are using survey data.

```
lin_mod1 <- svyglm(BMI ~ PhysActive*SmokeNow, design = df_adult)
lin_mod1 <- tidy(lin_mod1)

dif_n_smoking <- lin_mod1 %>%
  filter(term=="PhysActiveYes") %>%
  select(estimate)
dif_smoking <- lin_mod1 %>%
  filter(term%in%c("PhysActiveYes", "PhysActiveYes:SmokeNowYes")) %>%
  summarize(estimate = sum(estimate))

print(lin_mod1)
```

```
## # A tibble: 4 x 5
##   term                                estimate std.error statistic  p.value
```

```
##      <chr>                <dbl>      <dbl>      <dbl>      <dbl>
## 1 (Intercept)            30.5        0.210    146.    2.62e-44
## 2 PhysActiveYes          -2.35        0.236    -9.97  4.96e-11
## 3 SmokeNowYes            -2.24        0.267    -8.40  2.26e- 9
## 4 PhysActiveYes:SmokeNowYes  1.00        0.344     2.92  6.52e- 3
```

```
print(dif_smoking)
```

```
## # A tibble: 1 x 1
##   estimate
##   <dbl>
## 1    -1.35
```

```
print(dif_n_smoking)
```

```
## # A tibble: 1 x 1
##   estimate
##   <dbl>
## 1    -2.35
```

Above we have fitted a regression model where the relation of body mass index with physical activity varies by smoking status.

Conclusions and Recomendations

The relation between physical activity and smoking has a small p-value, suggesting that the relation does vary by smoking status. There is a major difference in the non smoker population where the difference between physically active and non-physically active people is large.

It would be beneficial to check any assumptions about our model, so we can conclude that physically active people tend to have lower body mass index, as do smokers. Although they have similar effect sizes, we probably wouldn't want to recommend smoking along with exercise.

In order to determine whether physical activity causes lower indexes, we need to use causal inference methods. Also, adjust and check for other possible factors for our model to give a more solid conclusion of the insights we covered.

```
lin_mod2 <- svyglm(BMI ~ PhysActive*SmokeNow + Race1 + Alcohol12PlusYr + Gender,
  design = df_adult)
tidy(lin_mod2)
```

```
## # A tibble: 10 x 5
##   term                estimate std.error statistic  p.value
##   <chr>                <dbl>      <dbl>      <dbl>    <dbl>
## 1 (Intercept)            33.2        0.316    105.    1.75e-33
## 2 PhysActiveYes          -2.11        0.273    -7.75  5.56e- 8
## 3 SmokeNowYes            -2.23        0.303    -7.34  1.40e- 7
## 4 Race1Hispanic          -1.47        0.420    -3.49  1.88e- 3
## 5 Race1Mexican           -0.191       0.464    -0.412 6.84e- 1
## 6 Race1White             -2.08        0.320    -6.49  1.04e- 6
## 7 Race1Other             -3.11        0.620    -5.01  4.09e- 5
## 8 Alcohol12PlusYrYes     -0.855       0.358    -2.39  2.50e- 2
## 9 Gendermale             -0.256       0.230    -1.11  2.78e- 1
## 10 PhysActiveYes:SmokeNowYes  0.737       0.387     1.90  6.92e- 2
```