My Analysis of the BMI dataset

Melinda Higgins

March 11, 2016

## My Report on the BMI toy dataset

This is my first Rmarkdown report written as of 2016-03-11 15:29:08.

The dataset we will be using is a simple data example. We downloaded this example from Melinda Higgins' github website at <https://github.com/melindahiggins2000/CDCRworkshopCode>.

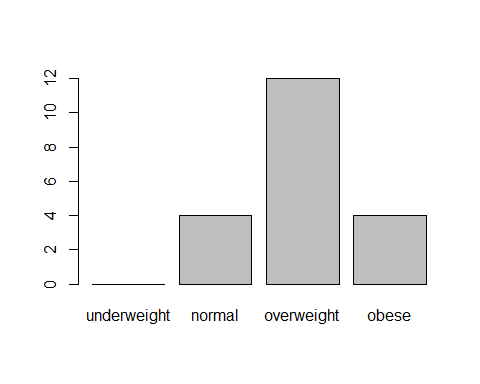
### Data Table

Here is the data:

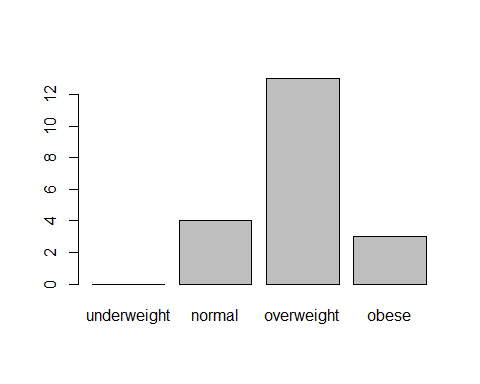
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Age | BMI PRE | BMI PRE Category | BMI POST | BMI POST Category | Gender |
| 45 | 23.35114 | normal | 22.57277 | normal | Male |
| 50 | 27.95901 | overweight | 27.79159 | overweight | Female |
| 35 | 22.26142 | normal | 21.01602 | normal | Female |
| 44 | 33.62564 | obese | 31.29052 | obese | Male |
| 32 | 32.95312 | obese | 30.24093 | obese | Male |
| 48 | 29.78997 | overweight | 26.17907 | overweight | Female |
| 50 | 22.94114 | normal | 22.94114 | normal | Male |
| 51 | 20.64644 | normal | 20.27105 | normal | Female |
| 46 | 26.95156 | overweight | 25.49908 | overweight | NA |
| 35 | 27.57341 | overweight | 29.02464 | overweight | Male |
| 36 | 29.21039 | overweight | 26.67035 | overweight | Male |
| 40 | 26.23996 | overweight | 25.58396 | overweight | Female |
| 45 | 25.24418 | overweight | 25.94541 | overweight | Female |
| 52 | 27.73176 | overweight | 25.42078 | overweight | Male |
| 24 | 29.79702 | overweight | 28.60514 | overweight | Male |
| 35 | 25.39656 | overweight | 25.25144 | overweight | Female |
| 51 | 27.06041 | overweight | 27.18341 | overweight | Male |
| 43 | 35.80508 | obese | 33.46996 | obese | Male |
| 36 | 28.54938 | overweight | 27.04678 | overweight | Female |
| 44 | 30.98891 | obese | 28.60514 | overweight | Male |

### Plots of BMI Categories

Here is a barplot of the BMI categories at PRE

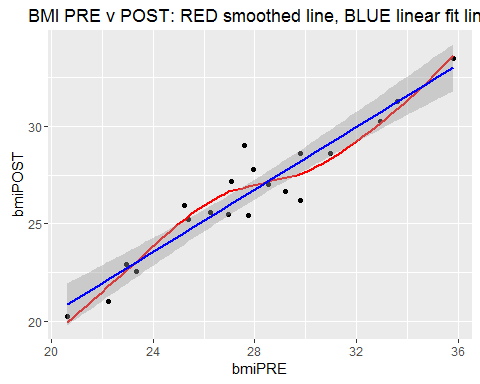


Here is a barplot of the BMI categories at PRE



### Comparisons of BMI PRE v POST

Here is a plot of the BMI PRE v POST. The linear model fit is shown with a **BLUE** line with 95% confidence intervals shown with grey shaded region and a **RED** line for the non-parametric *loess* smoothed line.



### Linear Model Results

The best fit linear model results were:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 4.3244779 | 1.7904717 | 2.415273 | 0.0265796 |
| bmiPRE | 0.8015493 | 0.0640344 | 12.517474 | 0.0000000 |

The ANOVA model fit for this model was:

kable(as.data.frame(anova(fit1)))

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Df | Sum Sq | Mean Sq | F value | Pr(>F) |
| bmiPRE | 1 | 184.00522 | 184.005223 | 156.6871 | 0 |
| Residuals | 18 | 21.13826 | 1.174348 | NA | NA |

### Compare by Gender

Here is the BMI PRE v POST paneled by Gender.

