Title: The war against fast food. Who's to blame for rising obesity rates of American children?
Team Name: 5 Minutes
Members: Tricia Chua, Ying-Hsi Liu, & Maria Medlarz

Abstract: The purpose of our paper was to see whether or not there was a correlation between the number of fast food restaurants and the rate of childhood obesity, per state. To test the significance of this correlation, we also tested for other factors including education, health spending per capita, and the length of roads in each state. We used publically available datasets that had relevant data for each state. From there, we regressed each coefficient against childhood obesity rates. We found that while the number of fast food restaurants had a linear relationship with the rate of childhood obesity, it was contradictory to what we were expecting--childhood obesity rates decreased as the number of fast food restaurants increased. In addition, when we compared the number of fast food restaurants with our other coefficients, we saw that it was not as significant. The rate of childhood obesity is affected by many different variables that this project was unable to fully capture and understand.
Acknowledgements: Panera. Our Professor, Tanya Rosenblat. Our Graduate Student Instructor, Yingzhi Liang. Our patient and forgiving grader, Shubham Rajora.

Cover Image from Fast Food USA,
http://cdn.livetolist.com/wp-content/uploads/2016/04/Fast-Food-USA.gif.

II. Introduction

In recent years, fast food franchises have begun to put effort toward adding healthier options not only to their overall menu, but to their children's menus as well. For example, McDonald's now offers yogurt and apple slices as side options with their Happy Meals and Burger King has added applesauce and fat free milk as side and drink choices with their Mighty Kids Meals. Based on these observations, it would seem that fast food is moving in a direction that would benefit America's children. However, despite seemingly healthier menu options arising, the children's menus are still populated by high calorie options like cheeseburgers, fries, and nuggets. We decided to further investigate the effect that the number of fast food restaurants in a given area have on that respective areas children's weight. In order to do so, we took datasets documenting childhood obesity rates by state from the Center for Disease Control and Prevention (CDC) and the number of fast food locations per state, controlling for population size.

Americans are often stereotyped as being fat or obese. We wanted to see how much truth there was in these stereotypes, specifically looking at American children. Our interest in the topic stemmed from our mutual interest in readily available food. As graduate students we are tight on time and being able to grab food quickly between classes can almost be considered essential to our busy schedules. A testament to our love of fast food is readily evident through the number of group meetings that we have held at the local Panera Bread. We believe that not only is this topic relevant to us, but also to our peers, and that they would be able to relate their own experiences to our research.

The majority of our research was based on one of the main concepts we used from the course, linear regression. We used linear regressions to determine whether or not our independent variable (fast food restaurants) could predict our dependent variable (childhood obesity). We also used dummy variables in order to divide the northern states from the southern states when comparing them to other outside variables. The outside variables we used were factors that may also be affecting childhood obesity in the United States.

When we completed our research we expected to find a correlation between the number of fast food restaurants in a given location and childhood obesity rates. We hypothesized that the more fast food restaurants there were, the higher the obesity rate would be among children. However, contrary to our beliefs, as obesity rates in children decreased, the number of fast food restaurants increased. In order to further examine the significance of our results, we decided to look at other data that may be able to explain childhood obesity rates.

Methodology

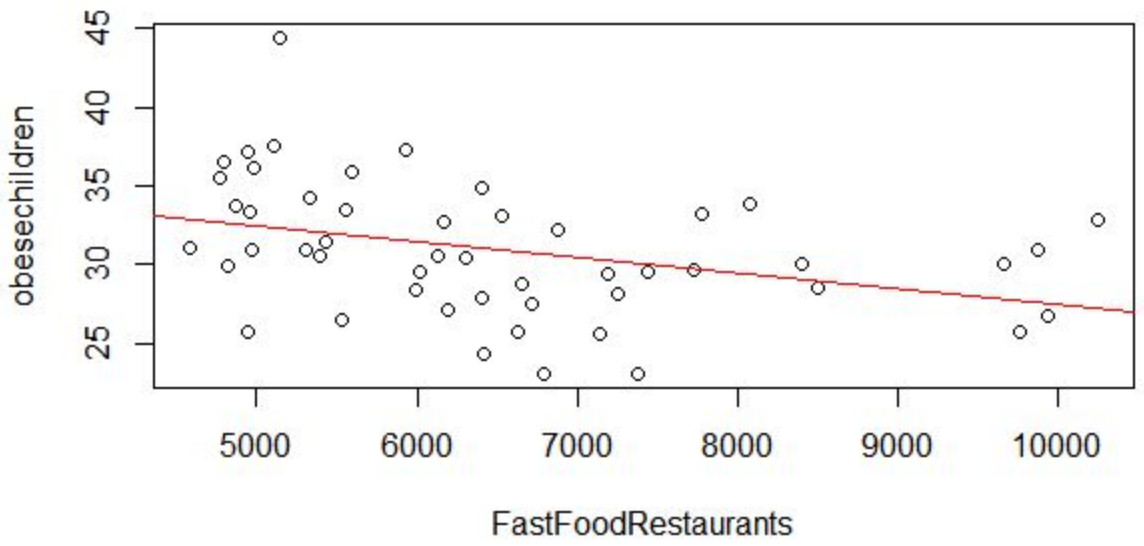
a. We used the following datasets for our project:
    i. Fast Food Maps (http://www.fastfoodmaps.com/data.html)
        1. To control for state population ((http://www.enchantedlearning.com/usa/states/population.shtml)
    ii. Prevalence of Childhood Obesity in the United States, 2011-2014 (https://www.cdc.gov/obesity/data/childhood.html)
    iii. Education (https://www.census.gov/compendia/statab/2012/tables/12s0233.xls, https://www.census.gov/prod/2012pubs/p20-566.pdf, https://www.census.gov/quickfacts/table/EDU635214/24)
    iv. Health Spending Per Capita by State (http://kff.org/other/state-indicator/health-spending-per-capita)
    v. Number of roads per state (http://blog.cubitplanning.com/2010/02/road-miles-by-state/)
        1. Controlled by land size per state (http://www.ipl.org/div/stateknow/popchart.html#statesbyland)
b. We used existing datasets we found online. While Google search was an effective first-step for finding much of our datasets, primarily the fast food maps, we also heavily relied on the University of Michigan's Library Guides.[1]
c. The advantages of using pre-existing datasets were that we did not have to set-up or conduct any surveys. The limitations, however, were that our findings and analyses were limited due to the the datasets we were able to find.
d. We analyzed our data primarily through RStudio. Our our full code is available in our Appendix.
   In addition to conducting standard regressions of our data using R, we also coded Southern versus non-Southern states to use as dummy variables.

Results

a. Based on your research question and data you collected, summarize your hypotheses
   Our null hypothesis is that the number of fast food locations in state has no effect on childhood obesity.
   Our alternative hypothesis is that the number of fast food locations has a positive effect on childhood obesity rates.
   We originally performed a regression on the number of fast food locations per state versus the percentage of childhood obesity. Our original graph did not appear linear and the p-values for childhood obesity were not significant. We took the population for each state and divided that by the number of fast food restaurants since we noticed larger population

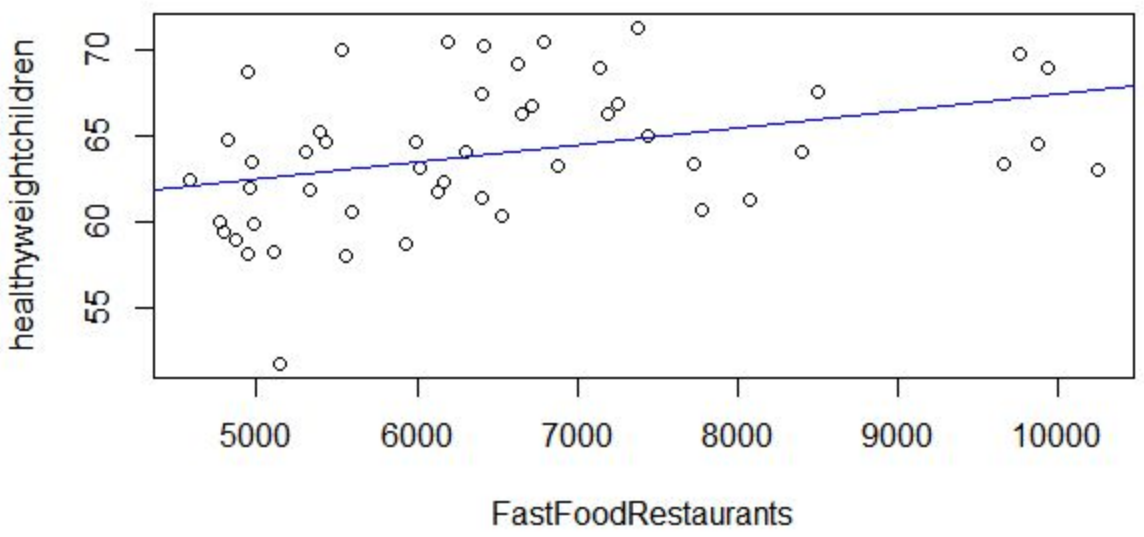
[1] http://guides.lib.umich.edu/globalstats, http://guides.lib.umich.edu/healthstats

states had many more fast food locations. After doing so, we performed another regression. This time the results were much more linear. But it seemed that the rate of obesity declined as the number of fast food restaurants increased.



We also performed a regression against the percentage of healthy as well as underweight kids per state. The results are presented below:

The percentage of underweight children seems unaffected by the number of fast food restaurants. There is a slightly positive but linear relation of the number of fast food restaurants and the percentage of healthy children as well. This appears to agree with our regression on obesity, which seemed to decrease with the number of fast food restaurants.



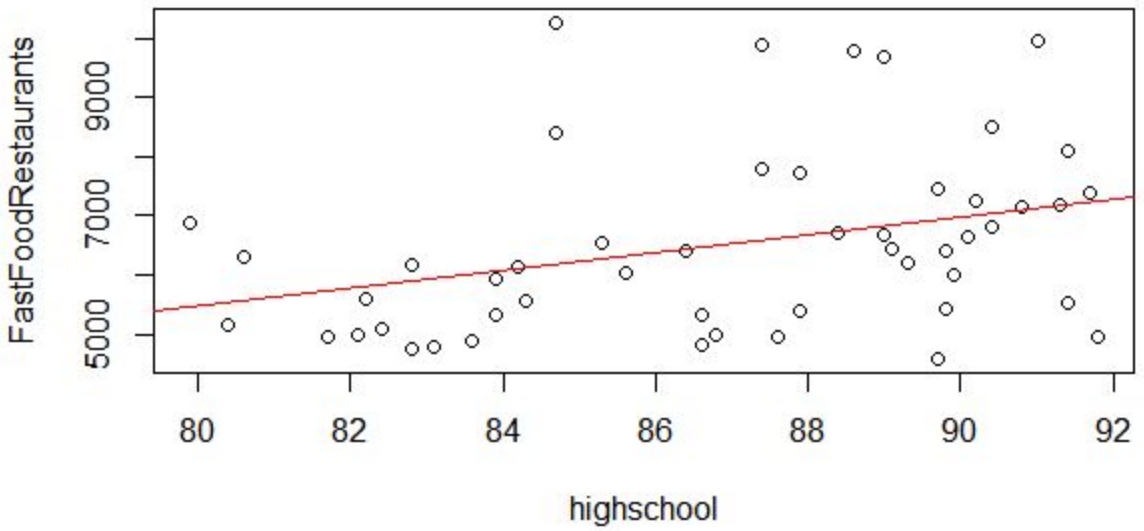
We performed some regressions to see which coefficients were the most significant

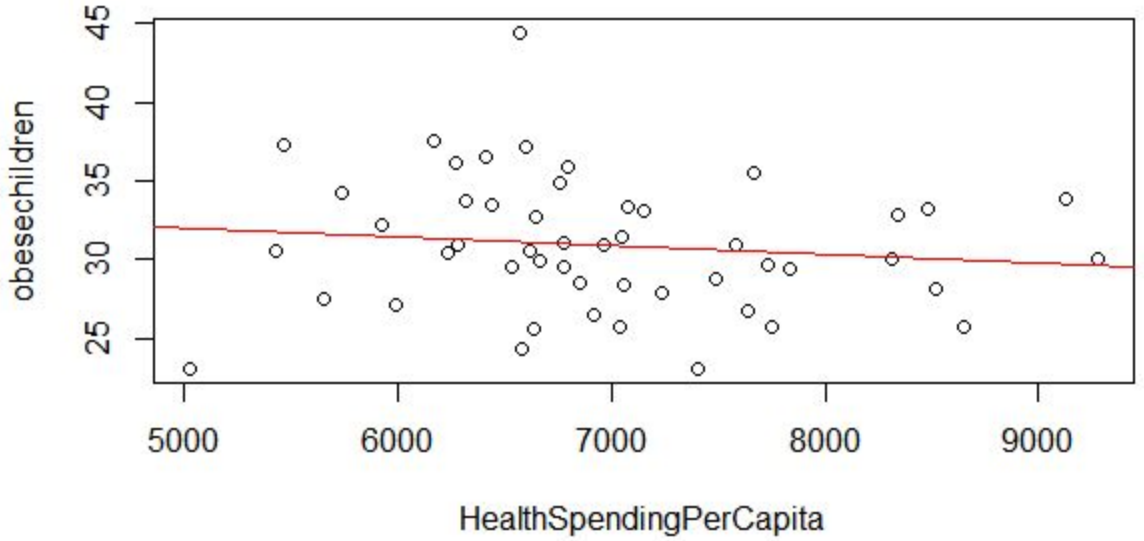
```
Residuals:
    Min      1Q  Median      3Q     Max
-4.4841 -1.5968 -0.0186  1.3529  6.5534

Coefficients:
                          Estimate Std. Error t value Pr(>|t|)
(Intercept)              1.029e+02  1.141e+01   9.020 1.79e-11 ***
bachelors               -2.895e-01  2.852e-01  -1.015   0.3158
highschool              -8.356e-01  1.721e-01  -4.855 1.63e-05 ***
advanceddegree           3.979e-01  5.013e-01   0.794   0.4317
HealthSpendingPerCapita  1.241e-03  6.214e-04   1.996   0.0523 .
FastFoodRestaurants     -6.247e-04  3.872e-04  -1.614   0.1139
RoadMiles                9.938e-03  1.871e-01   0.053   0.9579
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.595 on 43 degrees of freedom
Multiple R-squared:  0.6654,      Adjusted R-squared:  0.6187
F-statistic: 14.25 on 6 and 43 DF,  p-value: 7.325e-09
```

Based on this summary, we noticed that completing at least a high school degree and, to a smaller degree, Health Spending Per Capita, had much more significance on obesity rates than the number of Fast Food Restaurants.





We then perform a regression of rate of obesity vs number of fast food restaurants, Southern states or not and an interaction term of southern states or not and the number of fast food restaurants.

The dummy variable southern states is set to 1 in southern states, and 0 in not southern states. The southern states we define are Delaware,District of Columbia, Florida, Georgia

,Maryland, North Carolina, South Carolina, Virginia, West Virginia, Alabama, Kentucky , Mississippi, Tennessee, Arkansas, Louisiana, Oklahoma, and Texas.

```
> obesity<-read.csv("/Users/Sharon/Downloads/obesity_southern.csv")
> Y <- obesity$Overweight.or.obese..85th.percentile.or.above...
> X1 <- obesity$Number_restaurant
> X2 <- obesity$Southern
> X3 <- obesity$Number_restaurant*obesity$Southern
> lm_southern<-lm(Y~X1+X2+X3)
> summary(lm_southern)

Call:
lm(formula = Y ~ X1 + X2 + X3)

Residuals:
    Min      1Q  Median      3Q     Max
-6.0271 -2.2196  0.1175  1.9167  9.3352

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) 28.1872789  0.7179975  39.258  < 2e-16 ***
X1           0.0009492  0.0005040   1.883    0.066 .
X2           7.3245138  1.4528766   5.041 7.65e-06 ***
X3          -0.0017185  0.0009698  -1.772    0.083 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 3.21 on 46 degrees of freedom
Multiple R-squared:  0.4522,      Adjusted R-squared:  0.4165
F-statistic: 12.66 on 3 and 46 DF,  p-value: 3.69e-06
```

Based on the regression we ran, we were able to find out the estimated regression equation.

Rate of obesity = b0+ b1*Number of fast food restaurants + b2*southern or not
+b3*Number of fast food restaurants*southern or not
-->Rate of obesity = 28.187+0.0009b1+7.325b2 -0.0017b3

We can see that X2, the southern or not variable, is very significant. Therefore we can say that the children in southern states are more likely to be obese.
We then wanted to see how significant our dummy variables were after we added them with the rest of coefficients.

```
Residuals:
    Min      1Q  Median      3Q     Max
-3.8537 -1.6596 -0.3044  1.7765  6.2732
```

```
Coefficients:
                          Estimate Std. Error t value Pr(>|t|)
(Intercept)             89.6345567 14.6033827   6.138 2.75e-07 ***
X3                      -0.0008273  0.0008271  -1.000 0.323085
SouthernRestaurants      6.2965009  5.1259509   1.228 0.226317
bachelors               -0.2488554  0.2856539  -0.871 0.388730
highschool              -0.7010088  0.1970742  -3.557 0.000963 ***
advanceddegree           0.3253363  0.5012311   0.649 0.519910
HealthSpendingPerCapita  0.0011503  0.0006229   1.847 0.072022 .
FastFoodRestaurants     -0.0004266  0.0004062  -1.050 0.299788
RoadMiles                0.0063882  0.1864050   0.034 0.972828
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.583 on 41 degrees of freedom
Multiple R-squared:  0.6839,     Adjusted R-squared:  0.6222
F-statistic: 11.09 on 8 and 41 DF,  p-value: 3.616e-08
```

Here `X3` is the same as the previous regressions. While Southern states appear to be more likely to have higher rates of childhood obesity, the significance is much less compared to other coefficients.

b. The results were not as what we expected. Before running the regression, we supposed that as the number of fast food restaurants goes up in each state, the rate of obesity of children in that state will go up as well. We made this assumption because we think that as the state have more fast food restaurants per capita, the children will have more access to eating fast food, so they will more likely to be obese. However, the results are the opposite. The rate of obesity declined as the number of fast food restaurants increased. We were not sure that if the two variables really have a negative correlation or just coincidence, so we run regression on the number of fast food restaurants and the rate of underweight as well as the number of fast food restaurants and the rate of healthy weight. The regression on the number of fast food restaurants and the rate of healthy weight appears to have a positive correlation. Therefore, it seems safe to conclude that as the number of fast food restaurants per capita increases, the children are less likely to be obese in each state.

   We tried to find possible explanation for the results. First, we performed multiple regressions to see that if the factors such as level of education, health spending per capita, and the road miles in each state will affect the rate of obesity of children in each state. We found that the most significant variable is whether the rate of people having a high school degree. As more people having a high school degree in the state, the state is likely to have less obese children. The result can be explained as people with lower level of education(only

complete middle school or high school drop-out) are less likely to have health literacy, so they might not aware the obesity issues of their children. Also, people with lower level of education are more likely to have lower wages, so they can only choose foods that are not healthy and with low nutritions. Thus it makes the children obese.

We learn that people from southern states tend to have less healthy diet, so we try to find that if the location of the state will affect the obesity rate of children. We did a regression of rate of obesity vs number of fast food restaurants, Southern states or not and an interaction term of southern states or not and the number of fast food restaurants. For our results, we found that whether the state is a southern state or not is significantly correlated with the rate of obesity of the children in the state. We conclude that the rate of obesity of children might result from the daily diets people eat. Since southern states usually have sweeter foods or deep-fried foods, people having this kind of foods will easily get fat at a very young age. However, we also noticed that the significance of this correlation diminishes when other coefficients were added.

Conclusion

Summary of Findings

Our findings showed that as the number of fast food restaurants in a given location went up, the obesity rate in children in the corresponding area went down. This was found by performing multiple regressions after making sure that we took into account the size of a state when looking at how many restaurants were in that state. We also found that people from southern states have a higher obesity rate than those living in northern states. This was found by using dummy variables to separate the northern from the southern states when comparing them to different variables.

Limitations

A few limiting factors of our study can be attributed to data availability. Although we were able to narrow down obesity rates and number of fast food restaurants by state, we would have benefited from being able to use proportions of children in each area. We used the state population with the number of fast food restaurants in order to make different state data proportionate to each other, but with the population of children only we may have had more accurate results.

Additionally, we could have benefited from knowing more specific areas that the fast food restaurants were located. For example, if we had been able to narrow our data down to different towns within each state, we could have seen just how many children were nearby to various fast food restaurants. With the data we have now, our results could be based on restaurant locations that primarily serve adults.

Most importantly though, we all have very minimal knowledge of statistics. None of us have taken statistics or even math classes since our undergraduate careers, and do not use statistical methods on a daily basis. We did our best to use what we have learned this semester, but if we were more knowledgeable with R and other statistical programs, perhaps we could have gone further in depth with our studies. It is possible that there is an explanation for the negative correlation that we found between fast food restaurants and childhood obesity rates, but that the explanation will only be visible with more in depth regressions and methodology.

Further Research

If we had more time we would have liked to have compared our findings with various other topics. For instance, we have hypothesized that the mode of transportation that people use to commute and general layout of different cities could affect the childhood obesity rate. One in three people in the United States are overweight, and a key reason could be because of a lack of exercise during morning and afternoon commutes.[2] If a city is more accessible for biking to work and school, then the number of fast food restaurants may not have as much of an effect on childhood obesity rates because of a higher level of physical activity in the area.

We would have also liked to have looked into gym attendance in certain areas, as well as number of gym locations. Although it is unlikely that children would be spending time in gyms, we hypothesize that if parents are more active (going to the gym), then they would also encourage their children to be. The reason we would want to look into attendance, and not just membership, is because gym membership does not equal actual gym usage. Similar to our idea to compare obesity with gym membership numbers, we would have also liked to have compared the number of parents/adults that are overweight in each state. We suspect that parents that are overweight, and have unhealthy eating habits, would pass their unhealthy lifestyles to their children through example.

In addition to both of these further research topics, we think it would have been interesting to see how changes in legislation correlates with obesity rates in children. For example, when President Obama started his term as president, his wife, Michelle, became very involved in the "Let's Move" campaign, a campaign promoting healthy lifestyles in children.[3] She was often seen on TV commercials and other medias explaining the importance of fruits and vegetables in children's diets. "I am determined to work with folks across this country to change the way a generation of kids thinks about food and nutrition."[4] We would like to imagine that her efforts would have had a

[2] "Commute to work affects Food Choices and Obesity," *City Clock Magazine,* April 30, 2014, http://www.cityclock.org/commute-obesity-food-choices/#.WFAvETvwB2M.
[3] "Learn the Facts," Let's Move, The White House, http://www.letsmove.gov/learn-facts/epidemic-childhood-obesity.
[4] First Lady Michelle Obama, "About Let's Move: America's Move to Raise A Healthier Generation of Kids," Let's Move, The White House, http://www.letsmove.gov/about.

negative effect on childhood obesity rates, but then again we also thought that we would see different correlations then our project actually produced.

The most important question to be asked though, is which came first, childhood obesity or fast food restaurants? Perhaps childhood obesity has always been a problem, but in recent years the American public has turned fast food establishments into the scapegoat. It would definitely be an interesting study to pursue in the future to find out if these two variables really are related at all.

Benefits of doing this Research

When we first began with this study, we predicted that more fast food establishments would mean more obese children. After wrapping up our research though, we discovered just the opposite, at least in the correlations we ran. What was most interesting for us was the fact that our expectations were not met. We also had the opportunity to utilize our skills that we have learned throughout the semester, and apply them for practical purposes.

Appendix

Here are the other R Scripts that did not fit into the paper.

```
##FAST FOOD##
fastfood<-
read.csv("C:\\Users\\j\\Desktop\\fastfoodmaps_locations_2007.csv\\fastfoodmaps_location
s_2007.csv")
length(fastfood$state)
fastfoodedited <- read.csv("fastfoodedited.csv")


#trying to get number per state
FastFoodRestaurants <- fastfoodedited$Adjusted

##OBESITY##
obesity <- read.csv("obesity.csv")
obesechildren <- obesity$Overweight.or.obese..85th.percentile.or.above...

#healthy
healthyweightchildren <- obesity$Healthy.weight..5th.to.84th.percentile...
plot (FastFoodRestaurants, healthyweightchildren)
healthyweightchildren.lm <- lm(healthyweightchildren~FastFoodRestaurants)
abline(healthyweightchildren.lm, col="blue")
summary(healthyweightchildren.lm)

#underweight
underweightchildren <- obesity$Underweight..less.than.5th.percentile...
plot (FoodALL, underweightchildren)
underweightchildren.lm <- lm(underweightchildren~FastFoodRestaurants)
abline(underweightchildren.lm, col="blue")
summary(underweightchildren.lm)

#obese
plot(FastFoodRestaurants, obesechildren)
regression.lm <- lm(obesechildren~FastFoodRestaurants)
abline(regression.lm, col="red")
summary(regression.lm)


##Other Possibilities?#
#HealthSpendingPerState
HealthSpending <- read.csv("healthcare.csv")
HealthSpendingPerCapita <- HealthSpending$Health.Spending.per.Capita
plot(HealthSpendingPerCapita, obesechildren)
healthspending.lm <- lm(obesechildren~HealthSpendingPerCapita)
abline(healthspending.lm, col="red")
summary(healthspending.lm)
```

```
#roadmiles
Roads <- read.csv("roadmiles.csv")
RoadMiles <- Roads$Adjusted

#education
education <- read.csv("education.csv")
bachelors <- education$Bachelors
highschool <- education$High.School
advanceddegree <- education$Advanced

full.lm <- lm(obesechildren ~ bachelors + highschool + advanceddegree +
HealthSpendingPerCapita + FastFoodRestaurants + RoadMiles)
summary(full.lm)

plot(highschool, FastFoodRestaurants)
highschool.lm <- lm(FastFoodRestaurants ~ highschool)
abline(highschool.lm, col="red")
```