[Adult Consumption of Sugary Drinks (1+/day) - Number of Adults Reporting](http://a816-dohbesp.nyc.gov/IndicatorPublic/VisualizationData.aspx?id=2061,4466a0,96,Summarize" \l "summarize_collapse0)

**How Calculated:** Estimated number of adults who, on average reported having consumed one or more sugary drinks per day. Sugary drinks include soda, sweetened iced tea, sports drinks, fruit punch, and other fruit flavored drinks. (One drink equals 12-ounces.) Diet soda, sugar free drinks, 100% juice, and seltzer are not included.   
  
**Source(s):** New York City Community Health Survey (CHS)

[Adult Consumption of Sugary Drinks (1+/day) - Percent](http://a816-dohbesp.nyc.gov/IndicatorPublic/VisualizationData.aspx?id=2061,4466a0,96,Summarize#summarize_collapse1)

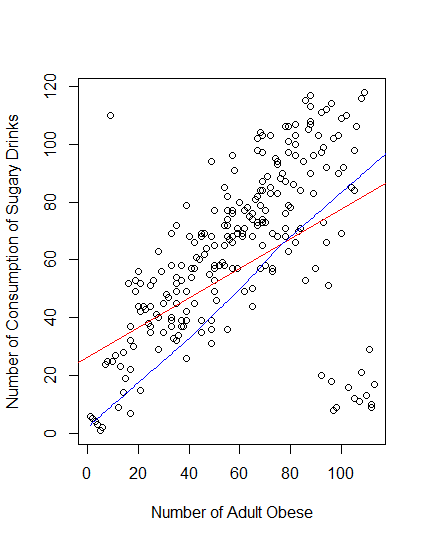
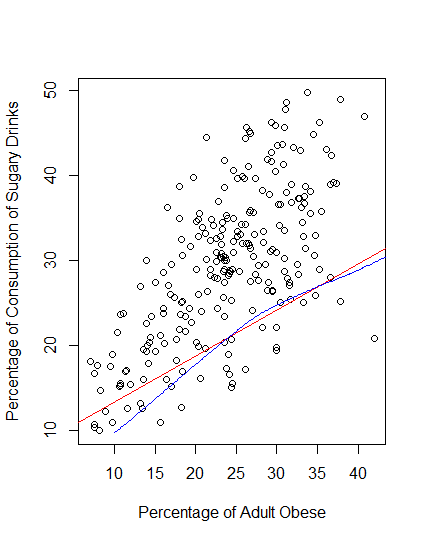
**How Calculated:** Estimated number of adults who, on average reported having consumed one or more sugary drinks per day, divided by all adults in the area; expressed as a percent. Sugary drinks include soda, sweetened iced tea, sports drinks, fruit punch, and other fruit flavored drinks. (One drink equals 12 ounces). Diet soda, sugar free drinks, 100% juice, and seltzer are not included.   
  
**Source(s):** New York City Community Health Survey (CHS)

[Overweight or Obese Adults - Number](http://a816-dohbesp.nyc.gov/IndicatorPublic/VisualizationData.aspx?id=2061,4466a0,96,Summarize#summarize_collapse0)

**How Calculated:** Estimated number of adults classified as overweight or obese; based on the Body Mass Index (BMI) calculated from self-reported weight and height, rounded to the nearest 1,000. A BMI between 25 and 29.9 is classified as overweight, and a BMI of 30 or greater is classified as obese.  
  
**Source(s):** New York City Community Health Survey (CHS)

[Overweight or Obese Adults - Percent](http://a816-dohbesp.nyc.gov/IndicatorPublic/VisualizationData.aspx?id=2061,4466a0,96,Summarize#summarize_collapse1)

**How Calculated:** Estimated number of adults classified as overweight or obese; based on the Body Mass Index (BMI) calculated from self-reported weight and height, expressed as a percent. A BMI between 25 and 29.9 is classified as overweight, and a BMI of 30 or greater is classified as obese.  
  
**Source(s):** New York City Community Health Survey (CHS)



**Number of people**

**Correlations**

[1] 0.52

Call:

lm(formula = Obese\_Num ~ Consum\_Num)

Residuals:

Min 1Q Median 3Q Max

-73.52 -15.61 -2.53 10.50 81.05

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 26.3523 3.7588 7.01 2.4e-11 \*\*\*

Consum\_Num 0.5107 0.0544 9.38 < 2e-16 \*\*\*

---

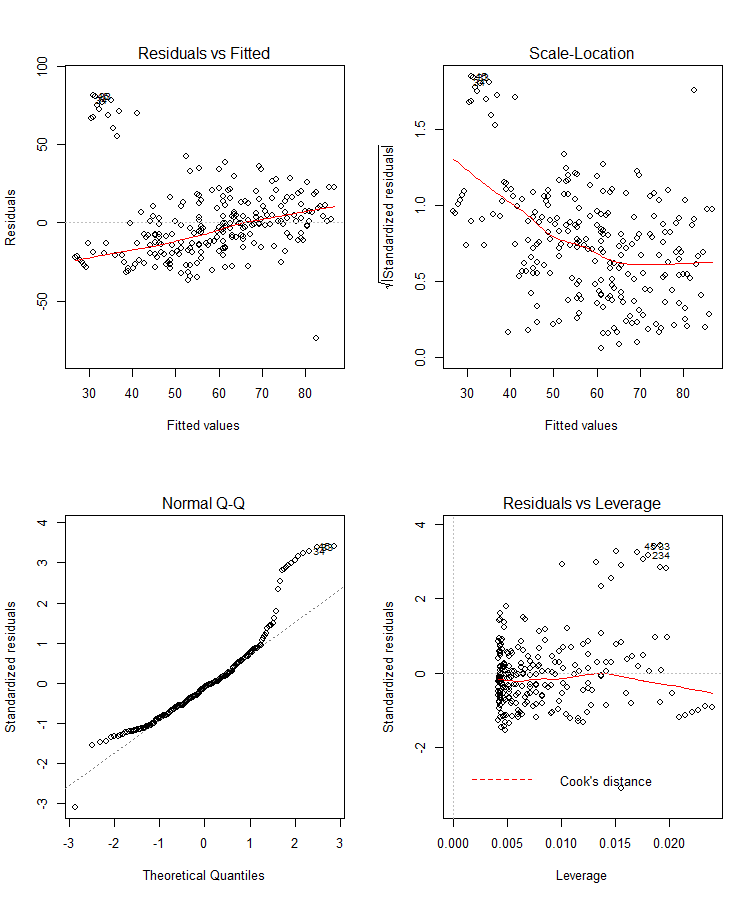
Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 24 on 238 degrees of freedom

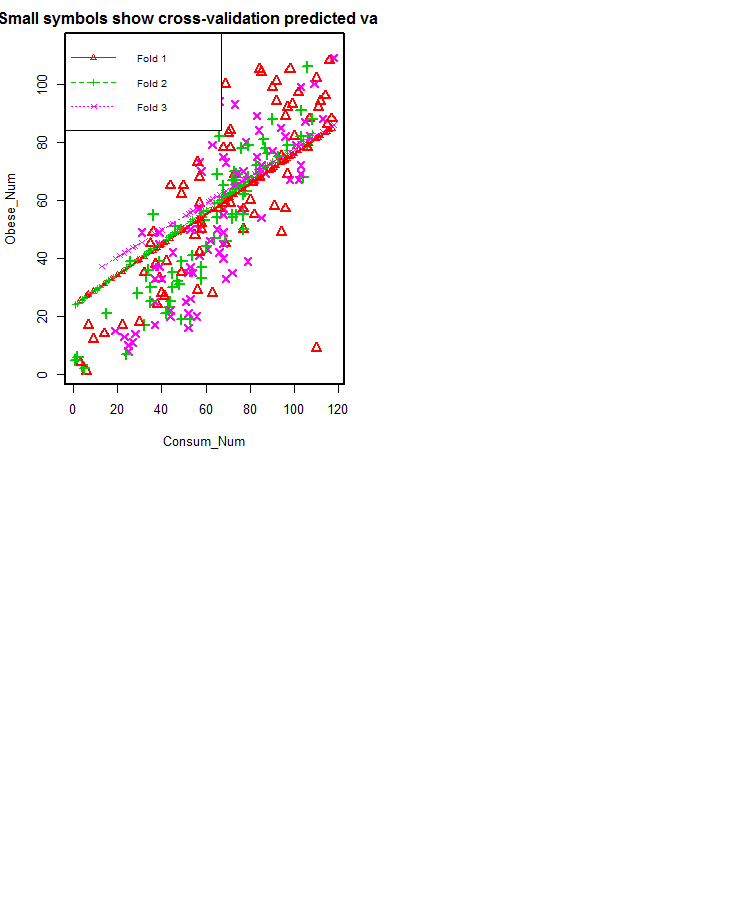
Multiple R-squared: 0.27, Adjusted R-squared: 0.267

F-statistic: 88 on 1 and 238 DF, p-value: <2e-16

**Equation: y=0.5107x+26.3523**



When conducting any statistical analysis it is important to evaluate how well the model fits the data and that the data meet the assumptions of the model.



**Cross-validation** is a technique used to protect against overfitting in a predictive model, particularly in a case where the amount of data may be limited. In cross-validation, you make a fixed number of folds (or partitions) of the data, run the analysis on each fold, and then average the overall error estimate.

Analysis of Variance Table

Response: Obese\_Num

Df Sum Sq Mean Sq F value Pr(>F)

Consum\_Num 1 50488 50488 88 <2e-16 \*\*\*

Residuals 238 136578 574

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

**Overall (Sum over all 80 folds)**

**ms**

**586**

**-> mean square error:**  representing the difference between the actual observations and the observation values predicted by the model, is used to determine the extent to which the model fits the data and whether the removal or some explanatory variables, simplifying the model, is possible without significantly harming the model's predictive ability.

**Percentage**

Correlations

[1] 0.6510899

Call:

lm(formula = Obese\_Per ~ Consum\_Per)

Residuals:

Min 1Q Median 3Q Max

-11.0651 -4.2132 -0.7027 3.9393 22.7861

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 7.93437 1.26036 6.295 1.46e-09 \*\*\*

Consum\_Per 0.54229 0.04098 13.234 < 2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

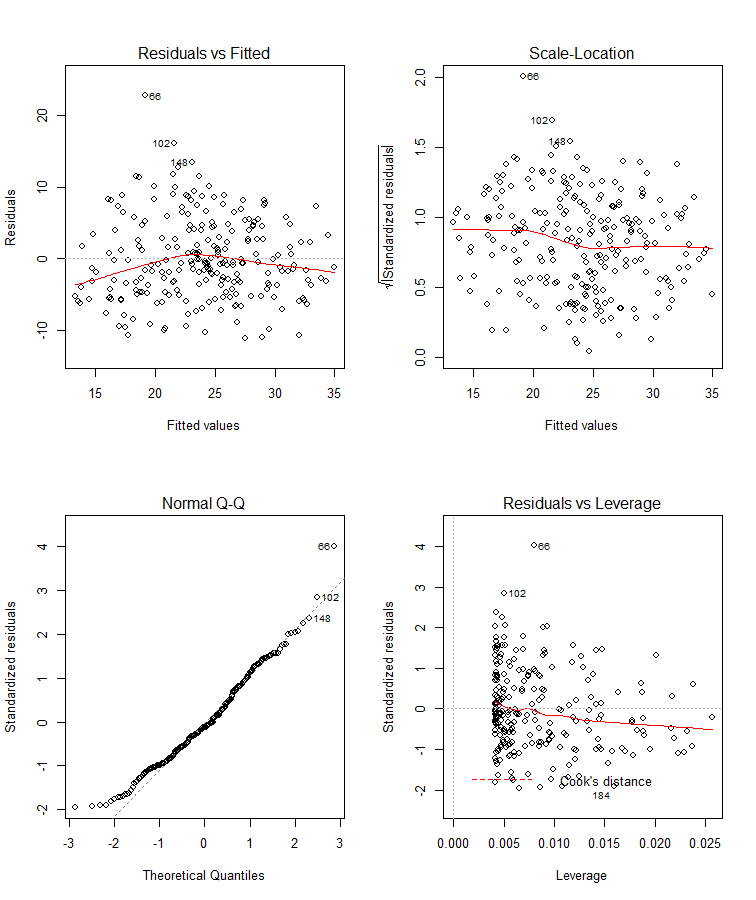
Residual standard error: 5.699 on 238 degrees of freedom

Multiple R-squared: 0.4239, Adjusted R-squared: 0.4215

F-statistic: 175.1 on 1 and 238 DF, p-value: < 2.2e-16

Model:

Y=0.54229x+7.93437



Analysis of Variance Table

Response: Obese\_Per

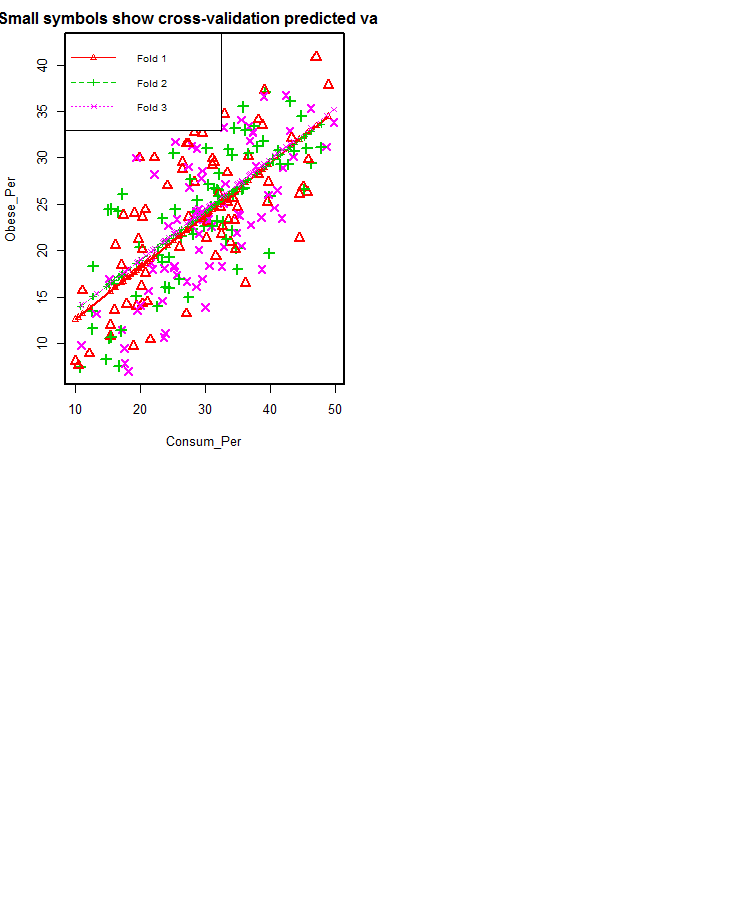
Df Sum Sq Mean Sq F value Pr(>F)

Consum\_Per 1 5688 5688 175 <2e-16 \*\*\*

Residuals 238 7730 32

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1



Overall (Sum over all 80 folds)

ms

32.6