STA305 tech report

Parth Brahmbhatt

March 29, 2021

### Importing the data

This is the data we will use, it contains all the stores with blocking as well as the nutritional information and its Price\_per\_Z\_mL.

data <- readxl::read\_excel("Data.xlsx", sheet = "All Stores with Stores Block")  
  
colnames(data) <- gsub(" ", "\_", colnames(data))  
colnames(data) <- gsub("\\(|\\)", "", colnames(data))

### Background

With the information presented below, we will check whether the data satisfies the assumptions of Guass-markov assumptions as well as the assumptions of linear regression.

### Scatterplot

## Walmart

# Price\_per\_Z\_mL vs Calories

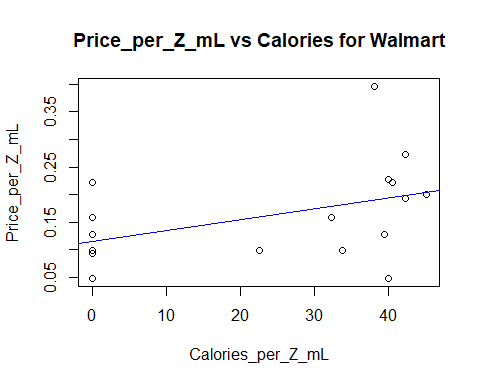
model <- lm(Price\_per\_Z\_mL ~ Calories\_per\_Z\_mL, data=subset(data, Store == 'Walmart'))  
  
summary(model)

##   
## Call:  
## lm(formula = Price\_per\_Z\_mL ~ Calories\_per\_Z\_mL, data = subset(data,   
## Store == "Walmart"))  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.14513 -0.03176 -0.01109 0.02979 0.20629   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.1152430 0.0252060 4.572 0.000236 \*\*\*  
## Calories\_per\_Z\_mL 0.0019596 0.0008875 2.208 0.040454 \*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.07673 on 18 degrees of freedom  
## Multiple R-squared: 0.2131, Adjusted R-squared: 0.1694   
## F-statistic: 4.875 on 1 and 18 DF, p-value: 0.04045

anova(model)

## Analysis of Variance Table  
##   
## Response: Price\_per\_Z\_mL  
## Df Sum Sq Mean Sq F value Pr(>F)   
## Calories\_per\_Z\_mL 1 0.028706 0.0287058 4.8755 0.04045 \*  
## Residuals 18 0.105980 0.0058878   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

plot(Price\_per\_Z\_mL ~ Calories\_per\_Z\_mL, main="Price\_per\_Z\_mL vs Calories for Walmart", data=subset(data, Store == 'Walmart'))  
  
abline(a = model$coefficients[1], b = model$coefficients[2], col="blue")



# Price\_per\_Z\_mL vs Fat (g)

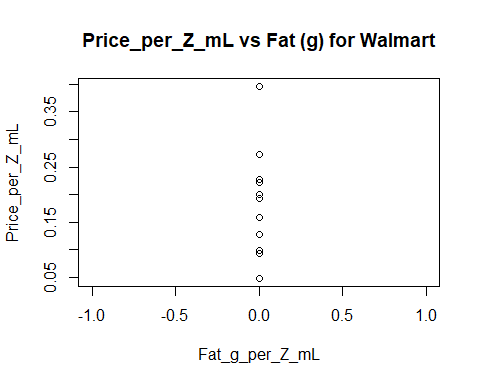
model <- lm(Price\_per\_Z\_mL ~ Fat\_g\_per\_Z\_mL, data=subset(data, Store == 'Walmart'))  
  
summary(model)

##   
## Call:  
## lm(formula = Price\_per\_Z\_mL ~ Fat\_g\_per\_Z\_mL, data = subset(data,   
## Store == "Walmart"))  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.10751 -0.05751 -0.02761 0.05007 0.23999   
##   
## Coefficients: (1 not defined because of singularities)  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.15601 0.01883 8.287 9.89e-08 \*\*\*  
## Fat\_g\_per\_Z\_mL NA NA NA NA   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.08419 on 19 degrees of freedom

anova(model)

## Analysis of Variance Table  
##   
## Response: Price\_per\_Z\_mL  
## Df Sum Sq Mean Sq F value Pr(>F)  
## Residuals 19 0.13469 0.0070887

plot(Price\_per\_Z\_mL ~ Fat\_g\_per\_Z\_mL, main="Price\_per\_Z\_mL vs Fat (g) for Walmart", data=subset(data, Store == 'Walmart'))



# Price\_per\_Z\_mL vs Sodium (mg)

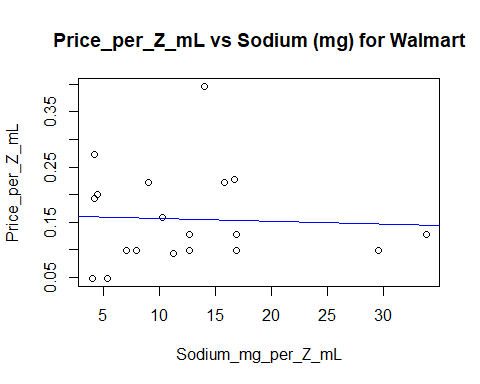
model <- lm(Price\_per\_Z\_mL ~ Sodium\_mg\_per\_Z\_mL, data=subset(data, Store == 'Walmart'))  
  
summary(model)

##   
## Call:  
## lm(formula = Price\_per\_Z\_mL ~ Sodium\_mg\_per\_Z\_mL, data = subset(data,   
## Store == "Walmart"))  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.11178 -0.05795 -0.02098 0.04664 0.24083   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.1623174 0.0362381 4.479 0.00029 \*\*\*  
## Sodium\_mg\_per\_Z\_mL -0.0005102 0.0024814 -0.206 0.83942   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.0864 on 18 degrees of freedom  
## Multiple R-squared: 0.002343, Adjusted R-squared: -0.05308   
## F-statistic: 0.04227 on 1 and 18 DF, p-value: 0.8394

anova(model)

## Analysis of Variance Table  
##   
## Response: Price\_per\_Z\_mL  
## Df Sum Sq Mean Sq F value Pr(>F)  
## Sodium\_mg\_per\_Z\_mL 1 0.000316 0.0003155 0.0423 0.8394  
## Residuals 18 0.134370 0.0074650

plot(Price\_per\_Z\_mL ~ Sodium\_mg\_per\_Z\_mL, main="Price\_per\_Z\_mL vs Sodium (mg) for Walmart", data=subset(data, Store == 'Walmart'))  
  
abline(a = model$coefficients[1], b = model$coefficients[2], col="blue")



# Price\_per\_Z\_mL vs Carbohydrates (g)

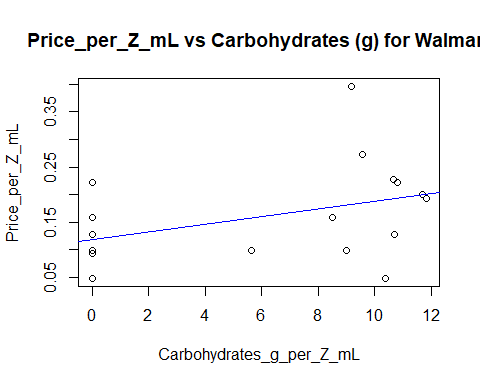
model <- lm(Price\_per\_Z\_mL ~ Carbohydrates\_g\_per\_Z\_mL, data=subset(data, Store == 'Walmart'))  
  
summary(model)

##   
## Call:  
## lm(formula = Price\_per\_Z\_mL ~ Carbohydrates\_g\_per\_Z\_mL, data = subset(data,   
## Store == "Walmart"))  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.14227 -0.03371 -0.01380 0.03083 0.21357   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.118435 0.025630 4.621 0.000212 \*\*\*  
## Carbohydrates\_g\_per\_Z\_mL 0.006956 0.003468 2.005 0.060184 .   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.0782 on 18 degrees of freedom  
## Multiple R-squared: 0.1826, Adjusted R-squared: 0.1372   
## F-statistic: 4.022 on 1 and 18 DF, p-value: 0.06018

anova(model)

## Analysis of Variance Table  
##   
## Response: Price\_per\_Z\_mL  
## Df Sum Sq Mean Sq F value Pr(>F)   
## Carbohydrates\_g\_per\_Z\_mL 1 0.024598 0.024598 4.0219 0.06018 .  
## Residuals 18 0.110088 0.006116   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

plot(Price\_per\_Z\_mL ~ Carbohydrates\_g\_per\_Z\_mL, main="Price\_per\_Z\_mL vs Carbohydrates (g) for Walmart", data=subset(data, Store == 'Walmart'))  
  
abline(a = model$coefficients[1], b = model$coefficients[2], col="blue")



# Price\_per\_Z\_mL vs Sugars (g)

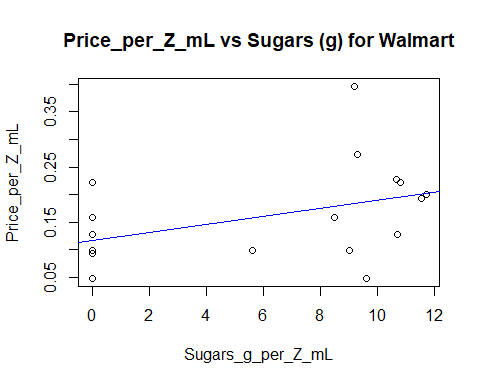
model <- lm(Price\_per\_Z\_mL ~ Sugars\_g\_per\_Z\_mL, data=subset(data, Store == 'Walmart'))  
  
summary(model)

##   
## Call:  
## lm(formula = Price\_per\_Z\_mL ~ Sugars\_g\_per\_Z\_mL, data = subset(data,   
## Store == "Walmart"))  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.13835 -0.03312 -0.01364 0.02887 0.21205   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.117456 0.025489 4.608 0.000218 \*\*\*  
## Sugars\_g\_per\_Z\_mL 0.007228 0.003494 2.069 0.053252 .   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.07775 on 18 degrees of freedom  
## Multiple R-squared: 0.1921, Adjusted R-squared: 0.1472   
## F-statistic: 4.279 on 1 and 18 DF, p-value: 0.05325

anova(model)

## Analysis of Variance Table  
##   
## Response: Price\_per\_Z\_mL  
## Df Sum Sq Mean Sq F value Pr(>F)   
## Sugars\_g\_per\_Z\_mL 1 0.02587 0.0258703 4.2794 0.05325 .  
## Residuals 18 0.10881 0.0060453   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

plot(Price\_per\_Z\_mL ~ Sugars\_g\_per\_Z\_mL, main="Price\_per\_Z\_mL vs Sugars (g) for Walmart", data=subset(data, Store == 'Walmart'))  
  
abline(a = model$coefficients[1], b = model$coefficients[2], col="blue")



# Price\_per\_Z\_mL vs Protein (g)

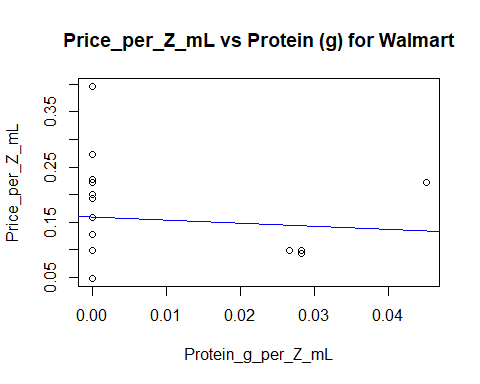
model <- lm(Price\_per\_Z\_mL ~ Protein\_g\_per\_Z\_mL, data=subset(data, Store == 'Walmart'))  
  
summary(model)

##   
## Call:  
## lm(formula = Price\_per\_Z\_mL ~ Protein\_g\_per\_Z\_mL, data = subset(data,   
## Store == "Walmart"))  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.11100 -0.05349 -0.03109 0.04659 0.23650   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.1595 0.0214 7.453 6.62e-07 \*\*\*  
## Protein\_g\_per\_Z\_mL -0.5438 1.4550 -0.374 0.713   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.08617 on 18 degrees of freedom  
## Multiple R-squared: 0.007701, Adjusted R-squared: -0.04743   
## F-statistic: 0.1397 on 1 and 18 DF, p-value: 0.7129

anova(model)

## Analysis of Variance Table  
##   
## Response: Price\_per\_Z\_mL  
## Df Sum Sq Mean Sq F value Pr(>F)  
## Protein\_g\_per\_Z\_mL 1 0.001037 0.0010372 0.1397 0.7129  
## Residuals 18 0.133649 0.0074249

plot(Price\_per\_Z\_mL ~ Protein\_g\_per\_Z\_mL, main="Price\_per\_Z\_mL vs Protein (g) for Walmart", data=subset(data, Store == 'Walmart'))  
  
abline(a = model$coefficients[1], b = model$coefficients[2], col="blue")



## Zehrs

# Price\_per\_Z\_mL vs Calories

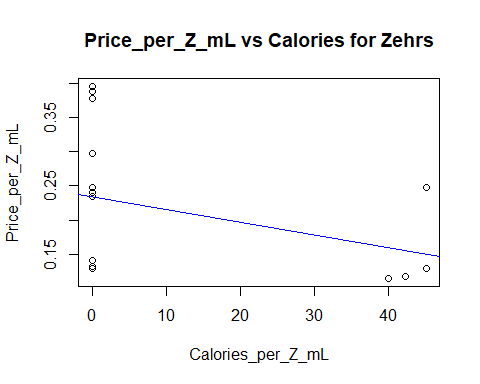
model <- lm(Price\_per\_Z\_mL ~ Calories\_per\_Z\_mL, data=subset(data, Store == 'Zehrs'))  
  
summary(model)

##   
## Call:  
## lm(formula = Price\_per\_Z\_mL ~ Calories\_per\_Z\_mL, data = subset(data,   
## Store == "Zehrs"))  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.10485 -0.09555 -0.01031 0.07154 0.16190   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.233723 0.025366 9.214 3.1e-08 \*\*\*  
## Calories\_per\_Z\_mL -0.001856 0.001315 -1.412 0.175   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.1015 on 18 degrees of freedom  
## Multiple R-squared: 0.09974, Adjusted R-squared: 0.04972   
## F-statistic: 1.994 on 1 and 18 DF, p-value: 0.175

anova(model)

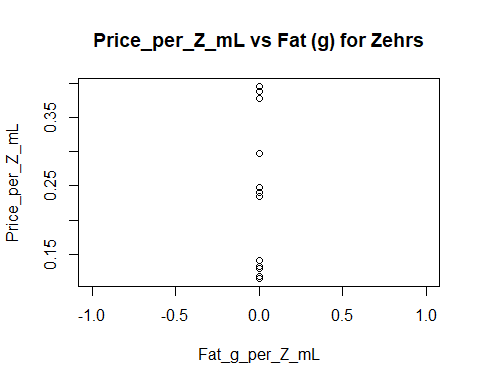
## Analysis of Variance Table  
##   
## Response: Price\_per\_Z\_mL  
## Df Sum Sq Mean Sq F value Pr(>F)  
## Calories\_per\_Z\_mL 1 0.020542 0.020542 1.9942 0.175  
## Residuals 18 0.185415 0.010301

plot(Price\_per\_Z\_mL ~ Calories\_per\_Z\_mL, main="Price\_per\_Z\_mL vs Calories for Zehrs", data=subset(data, Store == 'Zehrs'))  
  
abline(a = model$coefficients[1], b = model$coefficients[2], col="blue")



# Price\_per\_Z\_mL vs Fat (g)

model <- lm(Price\_per\_Z\_mL ~ Fat\_g\_per\_Z\_mL, data=subset(data, Store == 'Zehrs'))  
  
plot(Price\_per\_Z\_mL ~ Fat\_g\_per\_Z\_mL, main="Price\_per\_Z\_mL vs Fat (g) for Zehrs", data=subset(data, Store == 'Zehrs'))



# Price\_per\_Z\_mL vs Sodium (mg)

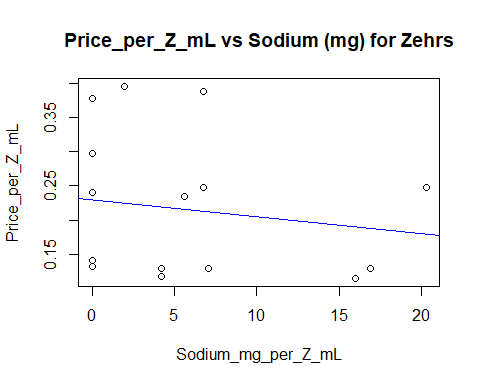
model <- lm(Price\_per\_Z\_mL ~ Sodium\_mg\_per\_Z\_mL, data=subset(data, Store == 'Zehrs'))  
  
summary(model)

##   
## Call:  
## lm(formula = Price\_per\_Z\_mL ~ Sodium\_mg\_per\_Z\_mL, data = subset(data,   
## Store == "Zehrs"))  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.10148 -0.08882 -0.02400 0.06781 0.17511   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.229007 0.029579 7.742 3.9e-07 \*\*\*  
## Sodium\_mg\_per\_Z\_mL -0.002458 0.003869 -0.635 0.533   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.1058 on 18 degrees of freedom  
## Multiple R-squared: 0.02194, Adjusted R-squared: -0.0324   
## F-statistic: 0.4037 on 1 and 18 DF, p-value: 0.5332

anova(model)

## Analysis of Variance Table  
##   
## Response: Price\_per\_Z\_mL  
## Df Sum Sq Mean Sq F value Pr(>F)  
## Sodium\_mg\_per\_Z\_mL 1 0.004518 0.0045182 0.4037 0.5332  
## Residuals 18 0.201438 0.0111910

plot(Price\_per\_Z\_mL ~ Sodium\_mg\_per\_Z\_mL, main="Price\_per\_Z\_mL vs Sodium (mg) for Zehrs", data=subset(data, Store == 'Zehrs'))  
  
abline(a = model$coefficients[1], b = model$coefficients[2], col="blue")



# Price\_per\_Z\_mL vs Carbohydrates (g)

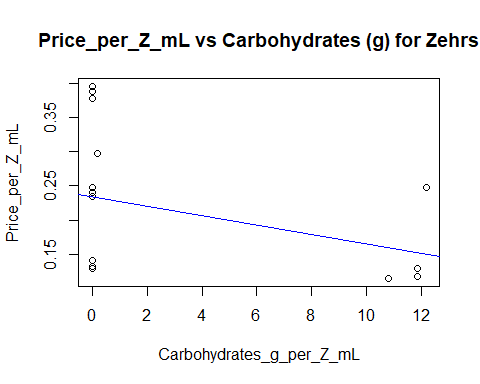
model <- lm(Price\_per\_Z\_mL ~ Carbohydrates\_g\_per\_Z\_mL, data=subset(data, Store == 'Zehrs'))  
  
summary(model)

##   
## Call:  
## lm(formula = Price\_per\_Z\_mL ~ Carbohydrates\_g\_per\_Z\_mL, data = subset(data,   
## Store == "Zehrs"))  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.10491 -0.09561 -0.01162 0.07246 0.16184   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.233788 0.025398 9.205 3.15e-08 \*\*\*  
## Carbohydrates\_g\_per\_Z\_mL -0.006862 0.004868 -1.410 0.176   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.1015 on 18 degrees of freedom  
## Multiple R-squared: 0.09942, Adjusted R-squared: 0.04939   
## F-statistic: 1.987 on 1 and 18 DF, p-value: 0.1757

anova(model)

## Analysis of Variance Table  
##   
## Response: Price\_per\_Z\_mL  
## Df Sum Sq Mean Sq F value Pr(>F)  
## Carbohydrates\_g\_per\_Z\_mL 1 0.020477 0.020477 1.9872 0.1757  
## Residuals 18 0.185480 0.010304

plot(Price\_per\_Z\_mL ~ Carbohydrates\_g\_per\_Z\_mL, main="Price\_per\_Z\_mL vs Carbohydrates (g) for Zehrs", data=subset(data, Store == 'Zehrs'))  
  
abline(a = model$coefficients[1], b = model$coefficients[2], col="blue")



# Price\_per\_Z\_mL vs Sugars (g)

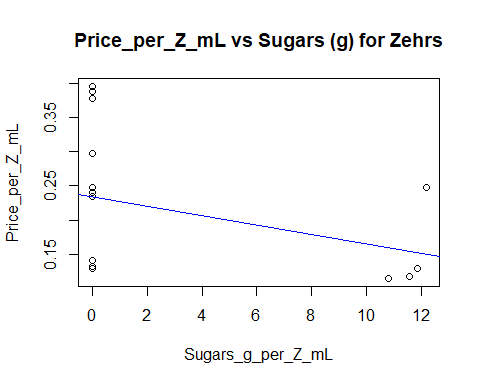
model <- lm(Price\_per\_Z\_mL ~ Sugars\_g\_per\_Z\_mL, data=subset(data, Store == 'Zehrs'))  
  
summary(model)

##   
## Call:  
## lm(formula = Price\_per\_Z\_mL ~ Sugars\_g\_per\_Z\_mL, data = subset(data,   
## Store == "Zehrs"))  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.10484 -0.09554 -0.01131 0.07163 0.16191   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.233717 0.025369 9.213 3.11e-08 \*\*\*  
## Sugars\_g\_per\_Z\_mL -0.006902 0.004892 -1.411 0.175   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.1015 on 18 degrees of freedom  
## Multiple R-squared: 0.09959, Adjusted R-squared: 0.04957   
## F-statistic: 1.991 on 1 and 18 DF, p-value: 0.1753

anova(model)

## Analysis of Variance Table  
##   
## Response: Price\_per\_Z\_mL  
## Df Sum Sq Mean Sq F value Pr(>F)  
## Sugars\_g\_per\_Z\_mL 1 0.020511 0.020511 1.9909 0.1753  
## Residuals 18 0.185445 0.010302

plot(Price\_per\_Z\_mL ~ Sugars\_g\_per\_Z\_mL, main="Price\_per\_Z\_mL vs Sugars (g) for Zehrs", data=subset(data, Store == 'Zehrs'))  
  
abline(a = model$coefficients[1], b = model$coefficients[2], col="blue")



# Price\_per\_Z\_mL vs Protein (g)

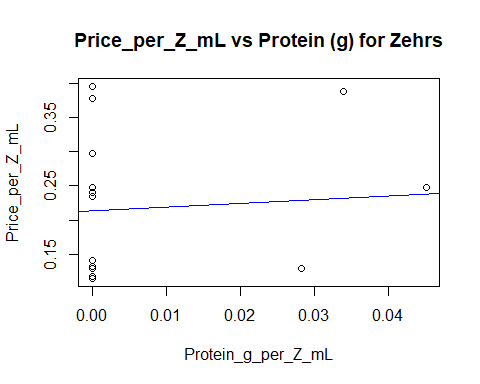
model <- lm(Price\_per\_Z\_mL ~ Protein\_g\_per\_Z\_mL, data=subset(data, Store == 'Zehrs'))  
  
summary(model)

##   
## Call:  
## lm(formula = Price\_per\_Z\_mL ~ Protein\_g\_per\_Z\_mL, data = subset(data,   
## Store == "Zehrs"))  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.10032 -0.08347 -0.03242 0.04535 0.18152   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.21410 0.02654 8.068 2.17e-07 \*\*\*  
## Protein\_g\_per\_Z\_mL 0.53574 1.71998 0.311 0.759   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.1067 on 18 degrees of freedom  
## Multiple R-squared: 0.005361, Adjusted R-squared: -0.0499   
## F-statistic: 0.09702 on 1 and 18 DF, p-value: 0.759

anova(model)

## Analysis of Variance Table  
##   
## Response: Price\_per\_Z\_mL  
## Df Sum Sq Mean Sq F value Pr(>F)  
## Protein\_g\_per\_Z\_mL 1 0.001104 0.0011042 0.097 0.759  
## Residuals 18 0.204852 0.0113807

plot(Price\_per\_Z\_mL ~ Protein\_g\_per\_Z\_mL, main="Price\_per\_Z\_mL vs Protein (g) for Zehrs", data=subset(data, Store == 'Zehrs'))  
  
abline(a = model$coefficients[1], b = model$coefficients[2], col="blue")



## Metro

# Price\_per\_Z\_mL vs Calories

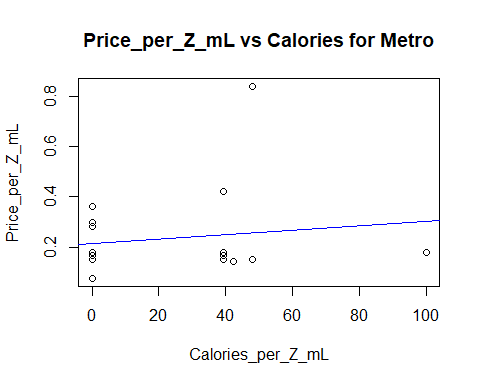
model <- lm(Price\_per\_Z\_mL ~ Calories\_per\_Z\_mL, data=subset(data, Store == 'Metro'))  
  
summary(model)

##   
## Call:  
## lm(formula = Price\_per\_Z\_mL ~ Calories\_per\_Z\_mL, data = subset(data,   
## Store == "Metro"))  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.13907 -0.08744 -0.06407 0.07223 0.58703   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.2135736 0.0488571 4.371 0.000368 \*\*\*  
## Calories\_per\_Z\_mL 0.0008698 0.0014061 0.619 0.543913   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.1703 on 18 degrees of freedom  
## Multiple R-squared: 0.02082, Adjusted R-squared: -0.03358   
## F-statistic: 0.3827 on 1 and 18 DF, p-value: 0.5439

anova(model)

## Analysis of Variance Table  
##   
## Response: Price\_per\_Z\_mL  
## Df Sum Sq Mean Sq F value Pr(>F)  
## Calories\_per\_Z\_mL 1 0.01110 0.011105 0.3827 0.5439  
## Residuals 18 0.52231 0.029017

plot(Price\_per\_Z\_mL ~ Calories\_per\_Z\_mL, main="Price\_per\_Z\_mL vs Calories for Metro", data=subset(data, Store == 'Metro'))  
  
abline(a = model$coefficients[1], b = model$coefficients[2], col="blue")



# Price\_per\_Z\_mL vs Fat (g)

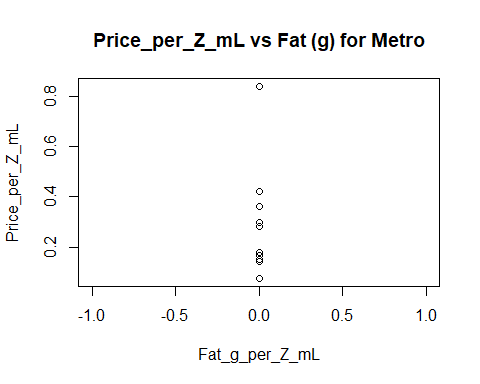
model <- lm(Price\_per\_Z\_mL ~ Fat\_g\_per\_Z\_mL, data=subset(data, Store == 'Metro'))  
  
summary(model)

##   
## Call:  
## lm(formula = Price\_per\_Z\_mL ~ Fat\_g\_per\_Z\_mL, data = subset(data,   
## Store == "Metro"))  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.15800 -0.08300 -0.06842 0.05330 0.60975   
##   
## Coefficients: (1 not defined because of singularities)  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.23250 0.03747 6.206 5.82e-06 \*\*\*  
## Fat\_g\_per\_Z\_mL NA NA NA NA   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.1676 on 19 degrees of freedom

anova(model)

## Analysis of Variance Table  
##   
## Response: Price\_per\_Z\_mL  
## Df Sum Sq Mean Sq F value Pr(>F)  
## Residuals 19 0.53342 0.028075

plot(Price\_per\_Z\_mL ~ Fat\_g\_per\_Z\_mL, main="Price\_per\_Z\_mL vs Fat (g) for Metro", data=subset(data, Store == 'Metro'))



# Price\_per\_Z\_mL vs Sodium (mg)

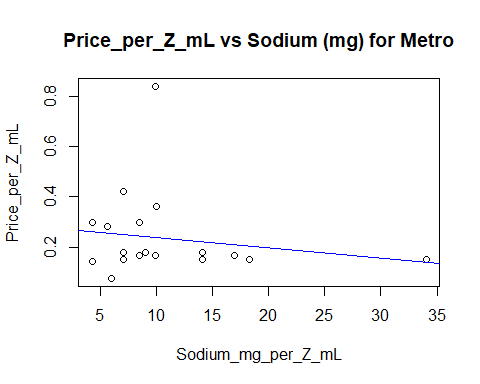
model <- lm(Price\_per\_Z\_mL ~ Sodium\_mg\_per\_Z\_mL, data=subset(data, Store == 'Metro'))  
  
summary(model)

##   
## Call:  
## lm(formula = Price\_per\_Z\_mL ~ Sodium\_mg\_per\_Z\_mL, data = subset(data,   
## Store == "Metro"))  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.17843 -0.07446 -0.05295 0.02993 0.60515   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.277531 0.072166 3.846 0.00118 \*\*  
## Sodium\_mg\_per\_Z\_mL -0.004101 0.005591 -0.733 0.47272   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.1696 on 18 degrees of freedom  
## Multiple R-squared: 0.02902, Adjusted R-squared: -0.02492   
## F-statistic: 0.538 on 1 and 18 DF, p-value: 0.4727

anova(model)

## Analysis of Variance Table  
##   
## Response: Price\_per\_Z\_mL  
## Df Sum Sq Mean Sq F value Pr(>F)  
## Sodium\_mg\_per\_Z\_mL 1 0.01548 0.015479 0.538 0.4727  
## Residuals 18 0.51794 0.028774

plot(Price\_per\_Z\_mL ~ Sodium\_mg\_per\_Z\_mL, main="Price\_per\_Z\_mL vs Sodium (mg) for Metro", data=subset(data, Store == 'Metro'))  
  
abline(a = model$coefficients[1], b = model$coefficients[2], col="blue")



# Price\_per\_Z\_mL vs Carbohydrates (g)

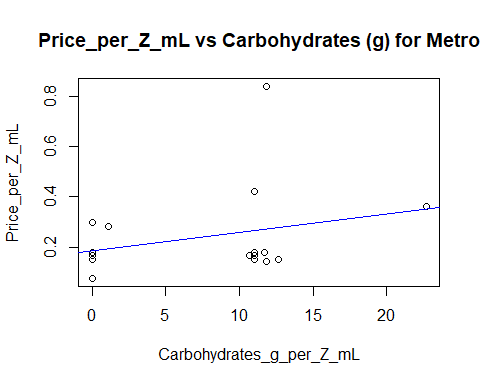
model <- lm(Price\_per\_Z\_mL ~ Carbohydrates\_g\_per\_Z\_mL, data=subset(data, Store == 'Metro'))  
  
summary(model)

##   
## Call:  
## lm(formula = Price\_per\_Z\_mL ~ Carbohydrates\_g\_per\_Z\_mL, data = subset(data,   
## Store == "Metro"))  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.13257 -0.10129 -0.03628 0.02939 0.56907   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.185779 0.050568 3.674 0.00174 \*\*  
## Carbohydrates\_g\_per\_Z\_mL 0.007388 0.005501 1.343 0.19596   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.1641 on 18 degrees of freedom  
## Multiple R-squared: 0.09108, Adjusted R-squared: 0.04058   
## F-statistic: 1.804 on 1 and 18 DF, p-value: 0.196

anova(model)

## Analysis of Variance Table  
##   
## Response: Price\_per\_Z\_mL  
## Df Sum Sq Mean Sq F value Pr(>F)  
## Carbohydrates\_g\_per\_Z\_mL 1 0.04858 0.048582 1.8037 0.196  
## Residuals 18 0.48484 0.026935

plot(Price\_per\_Z\_mL ~ Carbohydrates\_g\_per\_Z\_mL, main="Price\_per\_Z\_mL vs Carbohydrates (g) for Metro", data=subset(data, Store == 'Metro'))  
  
abline(a = model$coefficients[1], b = model$coefficients[2], col="blue")



# Price\_per\_Z\_mL vs Sugars (g)

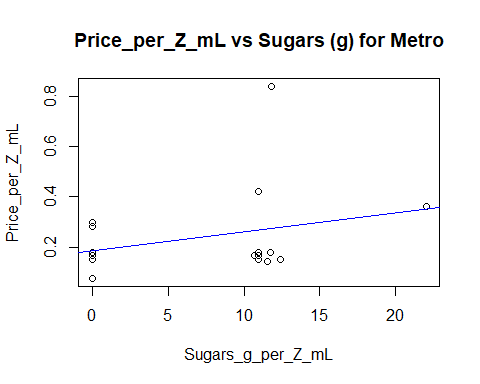
model <- lm(Price\_per\_Z\_mL ~ Sugars\_g\_per\_Z\_mL, data=subset(data, Store == 'Metro'))  
  
summary(model)

##   
## Call:  
## lm(formula = Price\_per\_Z\_mL ~ Sugars\_g\_per\_Z\_mL, data = subset(data,   
## Store == "Metro"))  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.13148 -0.10226 -0.03701 0.03407 0.56808   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.186513 0.050353 3.704 0.00162 \*\*  
## Sugars\_g\_per\_Z\_mL 0.007409 0.005551 1.335 0.19861   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.1642 on 18 degrees of freedom  
## Multiple R-squared: 0.09006, Adjusted R-squared: 0.0395   
## F-statistic: 1.781 on 1 and 18 DF, p-value: 0.1986

anova(model)

## Analysis of Variance Table  
##   
## Response: Price\_per\_Z\_mL  
## Df Sum Sq Mean Sq F value Pr(>F)  
## Sugars\_g\_per\_Z\_mL 1 0.04804 0.048038 1.7815 0.1986  
## Residuals 18 0.48538 0.026966

plot(Price\_per\_Z\_mL ~ Sugars\_g\_per\_Z\_mL, main="Price\_per\_Z\_mL vs Sugars (g) for Metro", data=subset(data, Store == 'Metro'))  
  
abline(a = model$coefficients[1], b = model$coefficients[2], col="blue")



# Price\_per\_Z\_mL vs Protein (g)

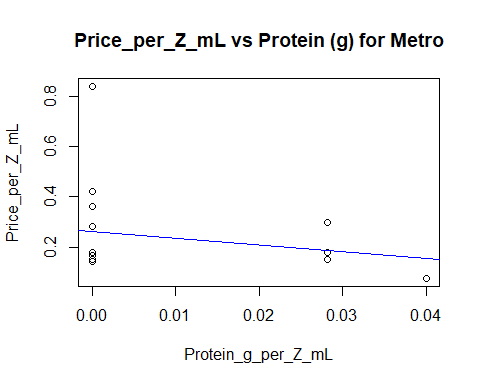
model <- lm(Price\_per\_Z\_mL ~ Protein\_g\_per\_Z\_mL, data=subset(data, Store == 'Metro'))  
  
summary(model)

##   
## Call:  
## lm(formula = Price\_per\_Z\_mL ~ Protein\_g\_per\_Z\_mL, data = subset(data,   
## Store == "Metro"))  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.11958 -0.09610 -0.05789 0.04157 0.58207   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.26019 0.04618 5.634 2.4e-05 \*\*\*  
## Protein\_g\_per\_Z\_mL -2.64903 2.58949 -1.023 0.32   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.1674 on 18 degrees of freedom  
## Multiple R-squared: 0.05495, Adjusted R-squared: 0.002442   
## F-statistic: 1.047 on 1 and 18 DF, p-value: 0.3199

anova(model)

## Analysis of Variance Table  
##   
## Response: Price\_per\_Z\_mL  
## Df Sum Sq Mean Sq F value Pr(>F)  
## Protein\_g\_per\_Z\_mL 1 0.02931 0.029309 1.0465 0.3199  
## Residuals 18 0.50411 0.028006

plot(Price\_per\_Z\_mL ~ Protein\_g\_per\_Z\_mL, main="Price\_per\_Z\_mL vs Protein (g) for Metro", data=subset(data, Store == 'Metro'))  
  
abline(a = model$coefficients[1], b = model$coefficients[2], col="blue")



## No Frills

# Price\_per\_Z\_mL vs Calories

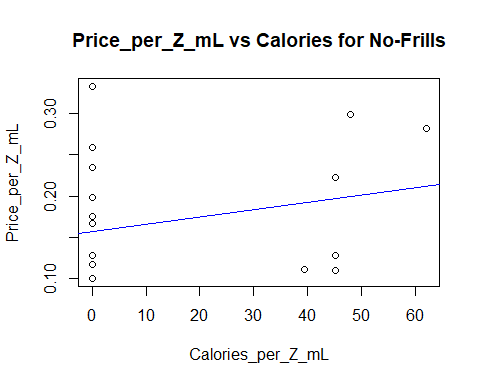
model <- lm(Price\_per\_Z\_mL ~ Calories\_per\_Z\_mL, data=subset(data, Store == 'No-Frills'))  
  
summary(model)

##   
## Call:  
## lm(formula = Price\_per\_Z\_mL ~ Calories\_per\_Z\_mL, data = subset(data,   
## Store == "No-Frills"))  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.08746 -0.05622 -0.03369 0.04926 0.17711   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.1562238 0.0202469 7.716 4.09e-07 \*\*\*  
## Calories\_per\_Z\_mL 0.0009066 0.0007710 1.176 0.255   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.0761 on 18 degrees of freedom  
## Multiple R-squared: 0.07134, Adjusted R-squared: 0.01974   
## F-statistic: 1.383 on 1 and 18 DF, p-value: 0.255

anova(model)

## Analysis of Variance Table  
##   
## Response: Price\_per\_Z\_mL  
## Df Sum Sq Mean Sq F value Pr(>F)  
## Calories\_per\_Z\_mL 1 0.008007 0.0080074 1.3827 0.255  
## Residuals 18 0.104243 0.0057913

plot(Price\_per\_Z\_mL ~ Calories\_per\_Z\_mL, main="Price\_per\_Z\_mL vs Calories for No-Frills", data=subset(data, Store == 'No-Frills'))  
  
abline(a = model$coefficients[1], b = model$coefficients[2], col="blue")



# Price\_per\_Z\_mL vs Fat (g)

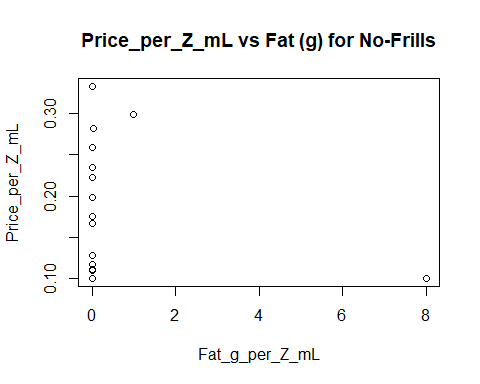
model <- lm(Price\_per\_Z\_mL ~ Fat\_g\_per\_Z\_mL, data=subset(data, Store == 'No-Frills'))  
  
summary(model)

##   
## Call:  
## lm(formula = Price\_per\_Z\_mL ~ Fat\_g\_per\_Z\_mL, data = subset(data,   
## Store == "No-Frills"))  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.07224 -0.06147 -0.03046 0.05368 0.16110   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.172236 0.018002 9.567 1.76e-08 \*\*\*  
## Fat\_g\_per\_Z\_mL -0.006893 0.009986 -0.690 0.499   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.07794 on 18 degrees of freedom  
## Multiple R-squared: 0.02579, Adjusted R-squared: -0.02834   
## F-statistic: 0.4765 on 1 and 18 DF, p-value: 0.4988

anova(model)

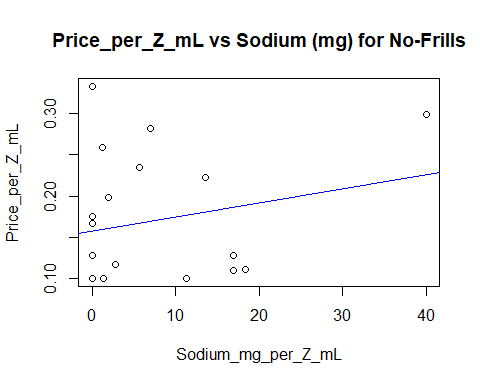
## Analysis of Variance Table  
##   
## Response: Price\_per\_Z\_mL  
## Df Sum Sq Mean Sq F value Pr(>F)  
## Fat\_g\_per\_Z\_mL 1 0.002895 0.0028947 0.4765 0.4988  
## Residuals 18 0.109355 0.0060753

plot(Price\_per\_Z\_mL ~ Fat\_g\_per\_Z\_mL, main="Price\_per\_Z\_mL vs Fat (g) for No-Frills", data=subset(data, Store == 'No-Frills'))



# Price\_per\_Z\_mL vs Sodium (mg)

model <- lm(Price\_per\_Z\_mL ~ Sodium\_mg\_per\_Z\_mL, data=subset(data, Store == 'No-Frills'))  
  
plot(Price\_per\_Z\_mL ~ Sodium\_mg\_per\_Z\_mL, main="Price\_per\_Z\_mL vs Sodium (mg) for No-Frills", data=subset(data, Store == 'No-Frills'))  
  
abline(a = model$coefficients[1], b = model$coefficients[2], col="blue")



# Price\_per\_Z\_mL vs Carbohydrates (g)

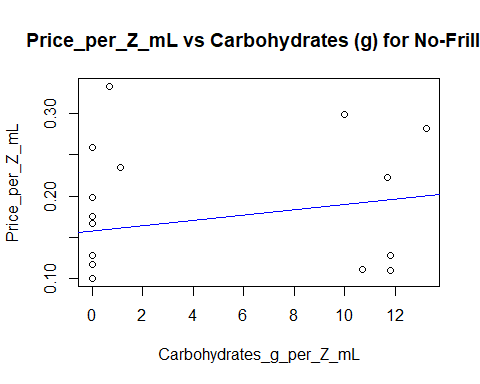
model <- lm(Price\_per\_Z\_mL ~ Carbohydrates\_g\_per\_Z\_mL, data=subset(data, Store == 'No-Frills'))  
  
summary(model)

##   
## Call:  
## lm(formula = Price\_per\_Z\_mL ~ Carbohydrates\_g\_per\_Z\_mL, data = subset(data,   
## Store == "No-Frills"))  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.08628 -0.05762 -0.03508 0.04926 0.17353   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.157618 0.020741 7.600 5.05e-07 \*\*\*  
## Carbohydrates\_g\_per\_Z\_mL 0.003236 0.003262 0.992 0.334   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.07689 on 18 degrees of freedom  
## Multiple R-squared: 0.05185, Adjusted R-squared: -0.0008287   
## F-statistic: 0.9843 on 1 and 18 DF, p-value: 0.3343

anova(model)

## Analysis of Variance Table  
##   
## Response: Price\_per\_Z\_mL  
## Df Sum Sq Mean Sq F value Pr(>F)  
## Carbohydrates\_g\_per\_Z\_mL 1 0.00582 0.0058198 0.9843 0.3343  
## Residuals 18 0.10643 0.0059128

plot(Price\_per\_Z\_mL ~ Carbohydrates\_g\_per\_Z\_mL, main="Price\_per\_Z\_mL vs Carbohydrates (g) for No-Frills", data=subset(data, Store == 'No-Frills'))  
  
abline(a = model$coefficients[1], b = model$coefficients[2], col="blue")



# Price\_per\_Z\_mL vs Sugars (g)

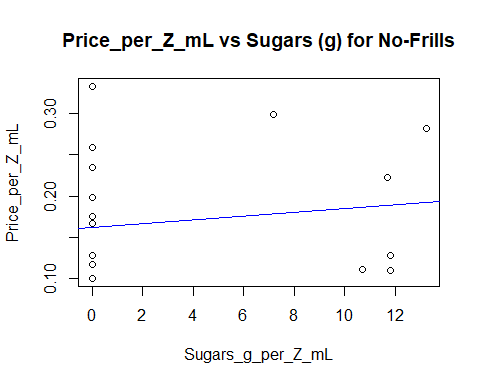
model <- lm(Price\_per\_Z\_mL ~ Sugars\_g\_per\_Z\_mL, data=subset(data, Store == 'No-Frills'))  
  
summary(model)

##   
## Call:  
## lm(formula = Price\_per\_Z\_mL ~ Sugars\_g\_per\_Z\_mL, data = subset(data,   
## Store == "No-Frills"))  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.07927 -0.06139 -0.03886 0.04640 0.17194   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.161393 0.020706 7.795 3.54e-07 \*\*\*  
## Sugars\_g\_per\_Z\_mL 0.002325 0.003362 0.692 0.498   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.07794 on 18 degrees of freedom  
## Multiple R-squared: 0.02588, Adjusted R-squared: -0.02823   
## F-statistic: 0.4783 on 1 and 18 DF, p-value: 0.498

anova(model)

## Analysis of Variance Table  
##   
## Response: Price\_per\_Z\_mL  
## Df Sum Sq Mean Sq F value Pr(>F)  
## Sugars\_g\_per\_Z\_mL 1 0.002905 0.0029054 0.4783 0.498  
## Residuals 18 0.109345 0.0060747

plot(Price\_per\_Z\_mL ~ Sugars\_g\_per\_Z\_mL, main="Price\_per\_Z\_mL vs Sugars (g) for No-Frills", data=subset(data, Store == 'No-Frills'))  
  
abline(a = model$coefficients[1], b = model$coefficients[2], col="blue")



# Price\_per\_Z\_mL vs Protein (g)

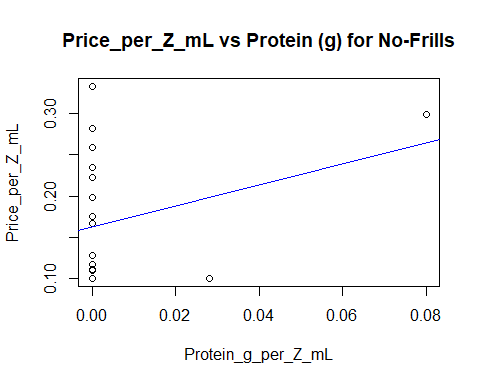
model <- lm(Price\_per\_Z\_mL ~ Protein\_g\_per\_Z\_mL, data=subset(data, Store == 'No-Frills'))  
  
summary(model)

##   
## Call:  
## lm(formula = Price\_per\_Z\_mL ~ Protein\_g\_per\_Z\_mL, data = subset(data,   
## Store == "No-Frills"))  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.09799 -0.05499 -0.03381 0.04264 0.17112   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.16221 0.01751 9.262 2.87e-08 \*\*\*  
## Protein\_g\_per\_Z\_mL 1.27858 0.92349 1.385 0.183   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.07507 on 18 degrees of freedom  
## Multiple R-squared: 0.09624, Adjusted R-squared: 0.04603   
## F-statistic: 1.917 on 1 and 18 DF, p-value: 0.1831

anova(model)

## Analysis of Variance Table  
##   
## Response: Price\_per\_Z\_mL  
## Df Sum Sq Mean Sq F value Pr(>F)  
## Protein\_g\_per\_Z\_mL 1 0.010803 0.0108033 1.9169 0.1831  
## Residuals 18 0.101447 0.0056359

plot(Price\_per\_Z\_mL ~ Protein\_g\_per\_Z\_mL, main="Price\_per\_Z\_mL vs Protein (g) for No-Frills", data=subset(data, Store == 'No-Frills'))  
  
abline(a = model$coefficients[1], b = model$coefficients[2], col="blue")

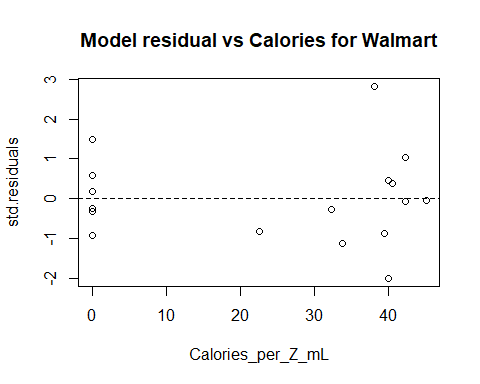


### Residual

## Walmart

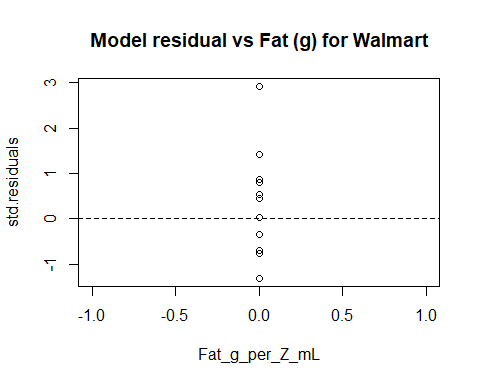
# Price\_per\_Z\_mL vs Calories

model <- lm(Price\_per\_Z\_mL ~ Calories\_per\_Z\_mL, data=subset(data, Store == 'Walmart'))  
  
res <- rstandard(model)  
  
plot(res ~ Calories\_per\_Z\_mL, main="Model residual vs Calories for Walmart", ylab="std.residuals", data=subset(data, Store == 'Walmart'))  
  
abline(h = 0, lty = 2)



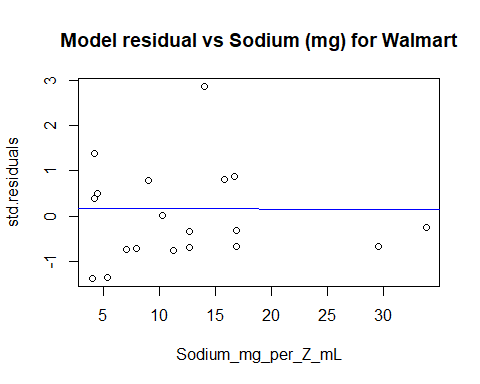
# Price\_per\_Z\_mL vs Fat (g)

model <- lm(Price\_per\_Z\_mL ~ Fat\_g\_per\_Z\_mL, data=subset(data, Store == 'Walmart'))  
  
res <- rstandard(model)  
  
plot(res ~ Fat\_g\_per\_Z\_mL, main="Model residual vs Fat (g) for Walmart", ylab="std.residuals", data=subset(data, Store == 'Walmart'))  
  
abline(h = 0, lty = 2)



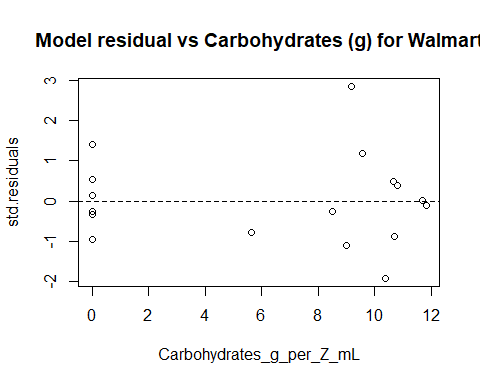
# Price\_per\_Z\_mL vs Sodium (mg)

model <- lm(Price\_per\_Z\_mL ~ Sodium\_mg\_per\_Z\_mL, data=subset(data, Store == 'Walmart'))  
  
res <- rstandard(model)  
  
plot(res ~ Sodium\_mg\_per\_Z\_mL, main="Model residual vs Sodium (mg) for Walmart",  
ylab="std.residuals", data=subset(data, Store == 'Walmart'))  
  
abline(a = model$coefficients[1], b = model$coefficients[2], col="blue")



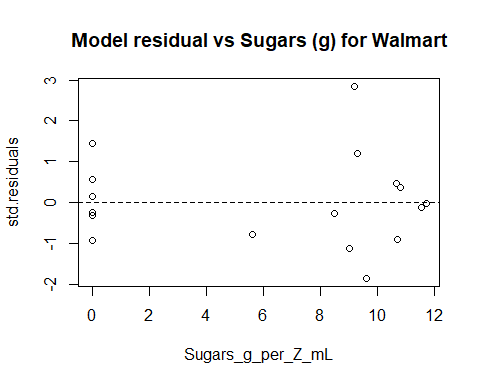
# Price\_per\_Z\_mL vs Carbohydrates (g)

model <- lm(Price\_per\_Z\_mL ~ Carbohydrates\_g\_per\_Z\_mL, data=subset(data, Store == 'Walmart'))  
  
res <- rstandard(model)  
  
plot(res ~ Carbohydrates\_g\_per\_Z\_mL, main="Model residual vs Carbohydrates (g) for Walmart", ylab="std.residuals", data=subset(data, Store == 'Walmart'))  
  
abline(h = 0, lty = 2)



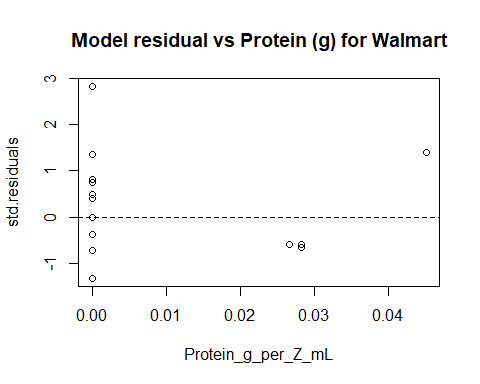
# Price\_per\_Z\_mL vs Sugars (g)

model <- lm(Price\_per\_Z\_mL ~ Sugars\_g\_per\_Z\_mL, data=subset(data, Store == 'Walmart'))  
  
res <- rstandard(model)  
  
plot(res ~ Sugars\_g\_per\_Z\_mL, main="Model residual vs Sugars (g) for Walmart", ylab="std.residuals", data=subset(data, Store == 'Walmart'))  
  
abline(h = 0, lty = 2)



# Price\_per\_Z\_mL vs Protein (g)

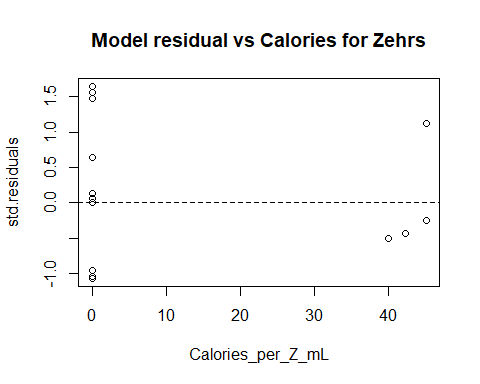
model <- lm(Price\_per\_Z\_mL ~ Protein\_g\_per\_Z\_mL, data=subset(data, Store == 'Walmart'))  
  
res <- rstandard(model)  
  
plot(res ~ Protein\_g\_per\_Z\_mL, ylab="std.residuals", main="Model residual vs Protein (g) for Walmart", data=subset(data, Store == 'Walmart'))  
  
abline(h = 0, lty = 2)



## Zehrs

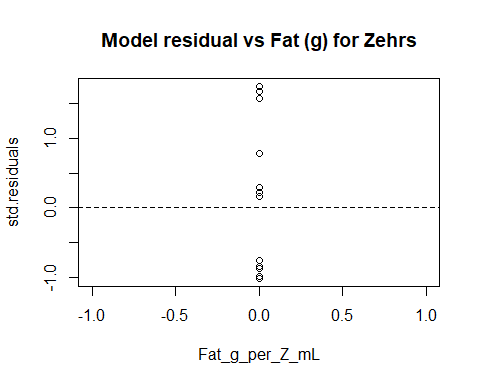
# Price\_per\_Z\_mL vs Calories

model <- lm(Price\_per\_Z\_mL ~ Calories\_per\_Z\_mL, data=subset(data, Store == 'Zehrs'))  
  
res <- rstandard(model)  
  
plot(res ~ Calories\_per\_Z\_mL, ylab="std.residuals", main="Model residual vs Calories for Zehrs", data=subset(data, Store == 'Zehrs'))  
  
abline(h = 0, lty = 2)



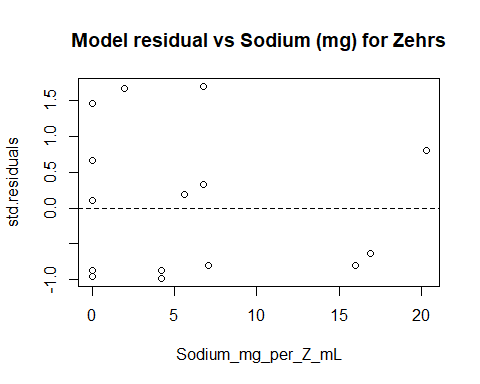
# Price\_per\_Z\_mL vs Fat (g)

model <- lm(Price\_per\_Z\_mL ~ Fat\_g\_per\_Z\_mL, data=subset(data, Store == 'Zehrs'))  
  
res <- rstandard(model)  
  
plot(res ~ Fat\_g\_per\_Z\_mL, ylab="std.residuals", main="Model residual vs Fat (g) for Zehrs", data=subset(data, Store == 'Zehrs'))  
  
abline(h = 0, lty = 2)



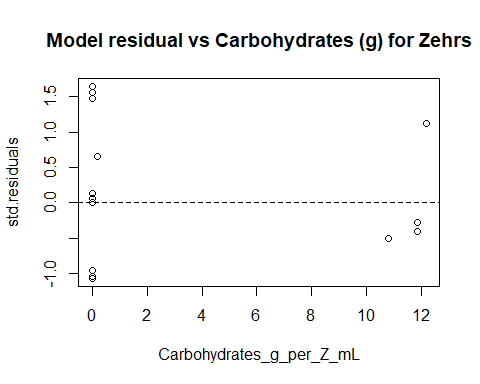
# Price\_per\_Z\_mL vs Sodium (mg)

model <- lm(Price\_per\_Z\_mL ~ Sodium\_mg\_per\_Z\_mL, data=subset(data, Store == 'Zehrs'))  
  
res <- rstandard(model)  
  
plot(res ~ Sodium\_mg\_per\_Z\_mL, ylab="std.residuals", main="Model residual vs Sodium (mg) for Zehrs", data=subset(data, Store == 'Zehrs'))  
  
abline(h = 0, lty = 2)



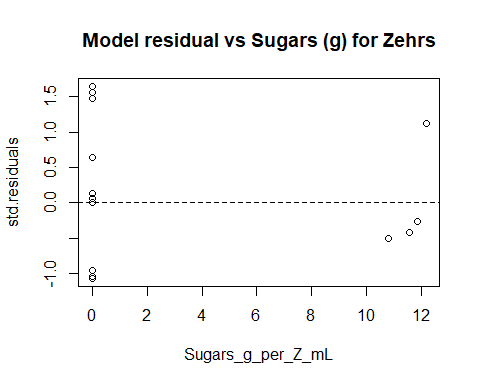
# Price\_per\_Z\_mL vs Carbohydrates (g)

model <- lm(Price\_per\_Z\_mL ~ Carbohydrates\_g\_per\_Z\_mL, data=subset(data, Store == 'Zehrs'))  
  
res <- rstandard(model)  
  
plot(res ~ Carbohydrates\_g\_per\_Z\_mL, ylab="std.residuals", main="Model residual vs Carbohydrates (g) for Zehrs", data=subset(data, Store == 'Zehrs'))  
  
abline(h = 0, lty = 2)



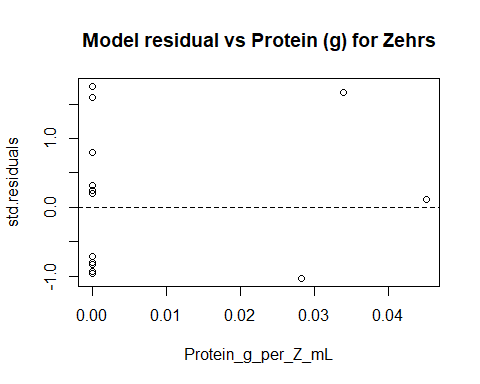
# Price\_per\_Z\_mL vs Sugars (g)

model <- lm(Price\_per\_Z\_mL ~ Sugars\_g\_per\_Z\_mL, data=subset(data, Store == 'Zehrs'))  
  
res <- rstandard(model)  
  
plot(res ~ Sugars\_g\_per\_Z\_mL, ylab="std.residuals", main="Model residual vs Sugars (g) for Zehrs", data=subset(data, Store == 'Zehrs'))  
  
abline(h = 0, lty = 2)



# Price\_per\_Z\_mL vs Protein (g)

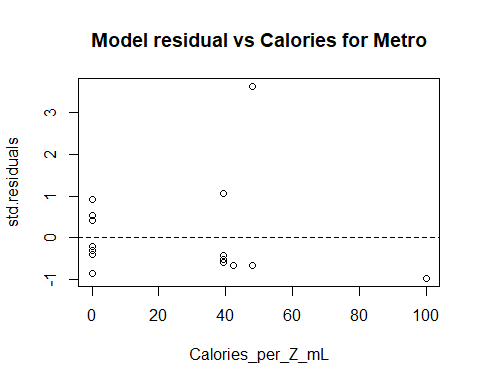
model <- lm(Price\_per\_Z\_mL ~ Protein\_g\_per\_Z\_mL, data=subset(data, Store == 'Zehrs'))  
  
res <- rstandard(model)  
  
plot(res ~ Protein\_g\_per\_Z\_mL, ylab="std.residuals", main="Model residual vs Protein (g) for Zehrs", data=subset(data, Store == 'Zehrs'))  
  
abline(h = 0, lty = 2)



## Metro

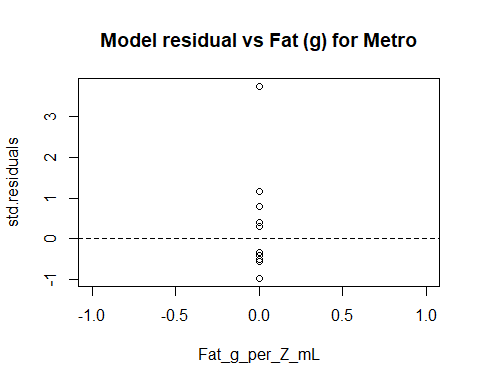
# Price\_per\_Z\_mL vs Calories

model <- lm(Price\_per\_Z\_mL ~ Calories\_per\_Z\_mL, data=subset(data, Store == 'Metro'))  
  
res <- rstandard(model)  
  
plot(res ~ Calories\_per\_Z\_mL, ylab="std.residuals", main="Model residual vs Calories for Metro", data=subset(data, Store == 'Metro'))  
  
abline(h = 0, lty = 2)



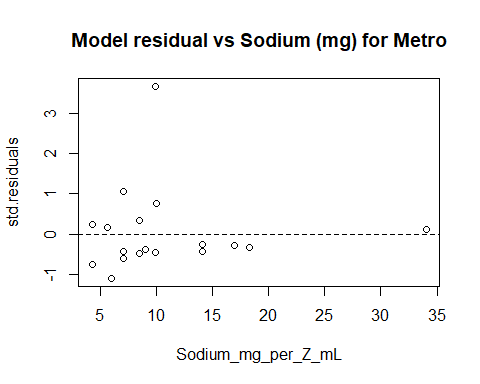
# Price\_per\_Z\_mL vs Fat (g)

model <- lm(Price\_per\_Z\_mL ~ Fat\_g\_per\_Z\_mL, data=subset(data, Store == 'Metro'))  
  
res <- rstandard(model)  
  
plot(res ~ Fat\_g\_per\_Z\_mL, ylab="std.residuals", main="Model residual vs Fat (g) for Metro", data=subset(data, Store == 'Metro'))  
  
abline(h = 0, lty = 2)



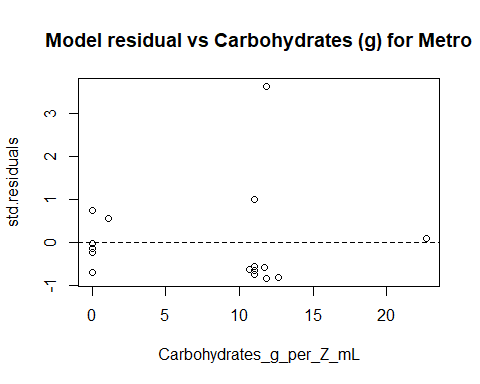
# Price\_per\_Z\_mL vs Sodium (mg)

model <- lm(Price\_per\_Z\_mL ~ Sodium\_mg\_per\_Z\_mL, data=subset(data, Store == 'Metro'))  
  
res <- rstandard(model)  
  
plot(res ~ Sodium\_mg\_per\_Z\_mL, ylab="std.residuals", main="Model residual vs Sodium (mg) for Metro", data=subset(data, Store == 'Metro'))  
  
abline(h = 0, lty = 2)



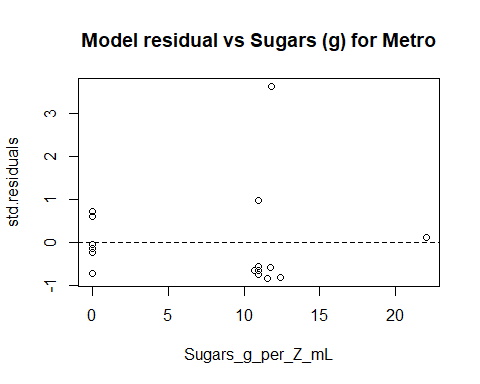
# Price\_per\_Z\_mL vs Carbohydrates (g)

model <- lm(Price\_per\_Z\_mL ~ Carbohydrates\_g\_per\_Z\_mL, data=subset(data, Store == 'Metro'))  
  
res <- rstandard(model)  
  
plot(res ~ Carbohydrates\_g\_per\_Z\_mL, ylab="std.residuals", main="Model residual vs Carbohydrates (g) for Metro", data=subset(data, Store == 'Metro'))  
  
abline(h = 0, lty = 2)



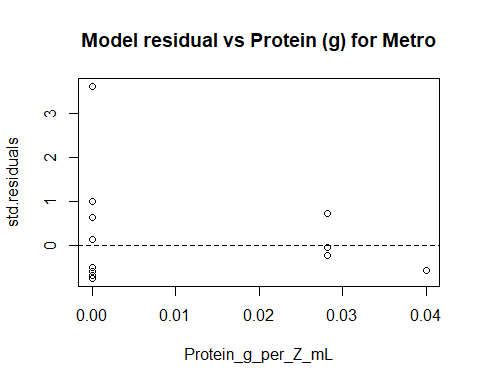
# Price\_per\_Z\_mL vs Sugars (g)

model <- lm(Price\_per\_Z\_mL ~ Sugars\_g\_per\_Z\_mL, data=subset(data, Store == 'Metro'))  
  
res <- rstandard(model)  
  
plot(res ~ Sugars\_g\_per\_Z\_mL, ylab="std.residuals", main="Model residual vs Sugars (g) for Metro", data=subset(data, Store == 'Metro'))  
  
abline(h = 0, lty = 2)



# Price\_per\_Z\_mL vs Protein (g)

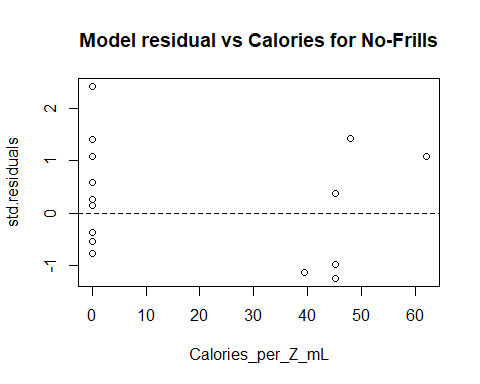
model <- lm(Price\_per\_Z\_mL ~ Protein\_g\_per\_Z\_mL, data=subset(data, Store == 'Metro'))  
  
res <- rstandard(model)  
  
plot(res ~ Protein\_g\_per\_Z\_mL, ylab="std.residuals", main="Model residual vs Protein (g) for Metro", data=subset(data, Store == 'Metro'))  
  
abline(h = 0, lty = 2)



## No-frills

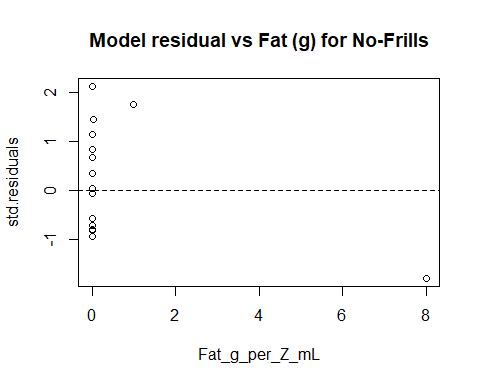
# Price\_per\_Z\_mL vs Calories

model <- lm(Price\_per\_Z\_mL ~ Calories\_per\_Z\_mL, data=subset(data, Store == 'No-Frills'))  
  
res <- rstandard(model)  
  
plot(res ~ Calories\_per\_Z\_mL, ylab="std.residuals", main="Model residual vs Calories for No-Frills", data=subset(data, Store == 'No-Frills'))  
  
abline(h = 0, lty = 2)



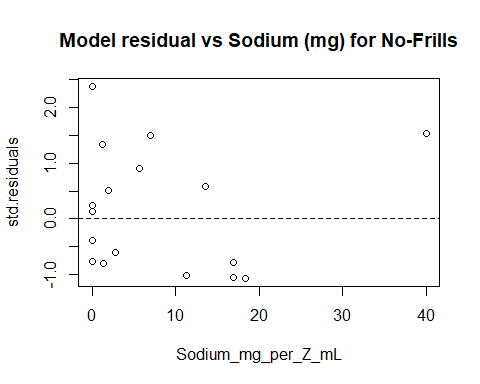
# Price\_per\_Z\_mL vs Fat (g)

model <- lm(Price\_per\_Z\_mL ~ Fat\_g\_per\_Z\_mL, data=subset(data, Store == 'No-Frills'))  
  
res <- rstandard(model)  
  
plot(res ~ Fat\_g\_per\_Z\_mL, ylab="std.residuals", main="Model residual vs Fat (g) for No-Frills", data=subset(data, Store == 'No-Frills'))  
  
abline(h = 0, lty = 2)



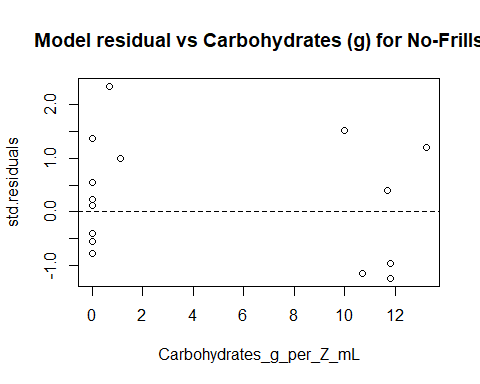
# Price\_per\_Z\_mL vs Sodium (mg)

model <- lm(Price\_per\_Z\_mL ~ Sodium\_mg\_per\_Z\_mL, data=subset(data, Store == 'No-Frills'))  
  
res <- rstandard(model)  
  
plot(res ~ Sodium\_mg\_per\_Z\_mL, ylab="std.residuals", main="Model residual vs Sodium (mg) for No-Frills", data=subset(data, Store == 'No-Frills'))  
  
abline(h = 0, lty = 2)



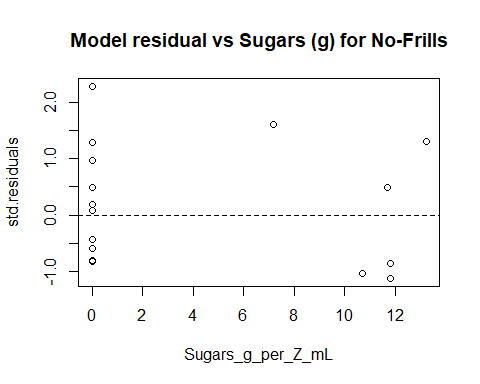
# Price\_per\_Z\_mL vs Carbohydrates (g)

model <- lm(Price\_per\_Z\_mL ~ Carbohydrates\_g\_per\_Z\_mL, data=subset(data, Store == 'No-Frills'))  
  
res <- rstandard(model)  
  
plot(res ~ Carbohydrates\_g\_per\_Z\_mL, ylab="std.residuals", main="Model residual vs Carbohydrates (g) for No-Frills", data=subset(data, Store == 'No-Frills'))  
  
abline(h = 0, lty = 2)



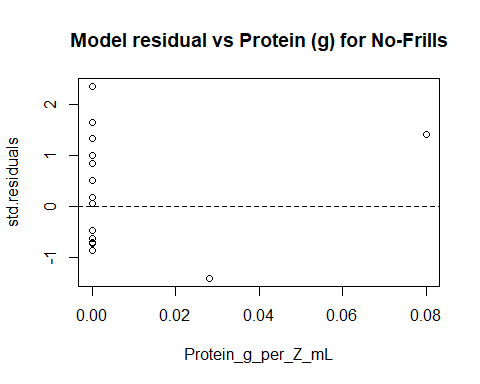
# Price\_per\_Z\_mL vs Sugars (g)

model <- lm(Price\_per\_Z\_mL ~ Sugars\_g\_per\_Z\_mL, data=subset(data, Store == 'No-Frills'))  
  
res <- rstandard(model)  
  
plot(res ~ Sugars\_g\_per\_Z\_mL, ylab="std.residuals", main="Model residual vs Sugars (g) for No-Frills", data=subset(data, Store == 'No-Frills'))  
  
abline(h = 0, lty = 2)



# Price\_per\_Z\_mL vs Protein (g)

model <- lm(Price\_per\_Z\_mL ~ Protein\_g\_per\_Z\_mL, data=subset(data, Store == 'No-Frills'))  
  
res <- rstandard(model)  
  
plot(res ~ Protein\_g\_per\_Z\_mL, ylab="std.residuals", main="Model residual vs Protein (g) for No-Frills", data=subset(data, Store == 'No-Frills'))  
  
abline(h = 0, lty = 2)

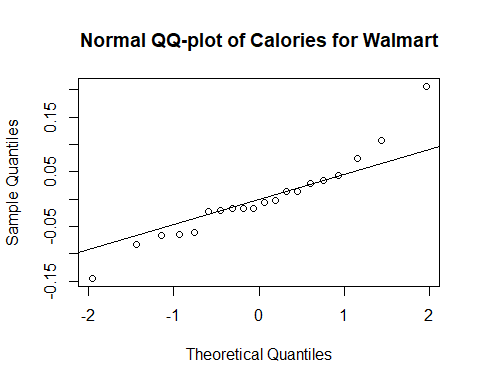


### QQplots

## Walmart

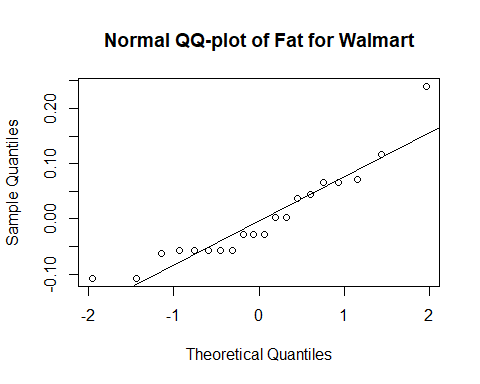
# Price\_per\_Z\_mL vs Calories

model <- lm(Price\_per\_Z\_mL ~ Calories\_per\_Z\_mL, data=subset(data, Store == 'Walmart'))  
  
qqnorm(residuals(model), main="Normal QQ-plot of Calories for Walmart")  
qqline(residuals(model))



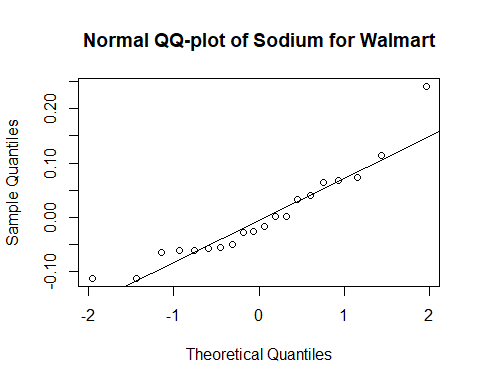
# Price\_per\_Z\_mL vs Fat (g)

model <- lm(Price\_per\_Z\_mL ~ Fat\_g\_per\_Z\_mL, data=subset(data, Store == 'Walmart'))  
  
qqnorm(residuals(model), main="Normal QQ-plot of Fat for Walmart")  
qqline(residuals(model))



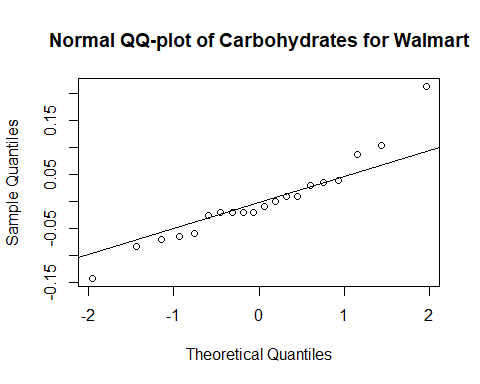
# Price\_per\_Z\_mL vs Sodium (mg)

model <- lm(Price\_per\_Z\_mL ~ Sodium\_mg\_per\_Z\_mL, data=subset(data, Store == 'Walmart'))  
  
qqnorm(residuals(model), main="Normal QQ-plot of Sodium for Walmart")  
qqline(residuals(model))



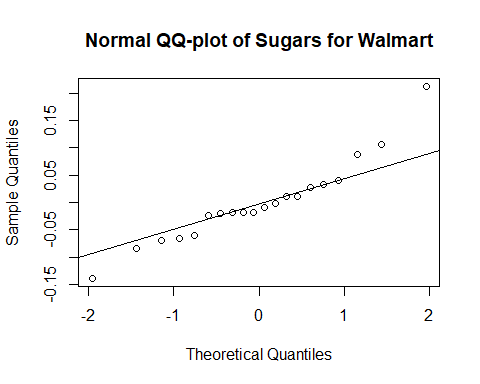
# Price\_per\_Z\_mL vs Carbohydrates (g)

model <- lm(Price\_per\_Z\_mL ~ Carbohydrates\_g\_per\_Z\_mL, data=subset(data, Store == 'Walmart'))  
  
qqnorm(residuals(model), main="Normal QQ-plot of Carbohydrates for Walmart")  
qqline(residuals(model))



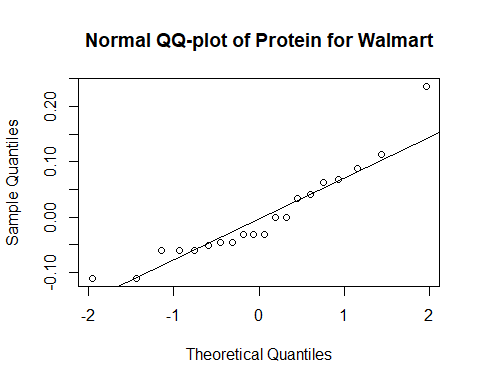
# Price\_per\_Z\_mL vs Sugars (g)

model <- lm(Price\_per\_Z\_mL ~ Sugars\_g\_per\_Z\_mL, data=subset(data, Store == 'Walmart'))  
  
qqnorm(residuals(model), main="Normal QQ-plot of Sugars for Walmart")  
qqline(residuals(model))



# Price\_per\_Z\_mL vs Protein (g)

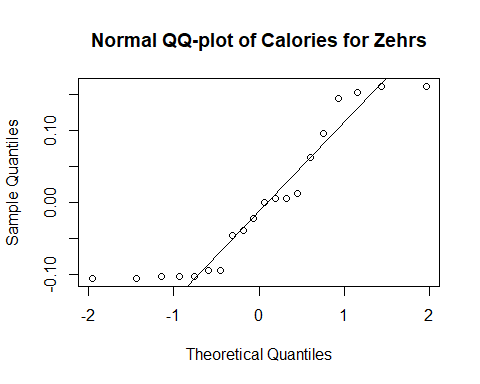
model <- lm(Price\_per\_Z\_mL ~ Protein\_g\_per\_Z\_mL, data=subset(data, Store == 'Walmart'))  
  
qqnorm(residuals(model), main="Normal QQ-plot of Protein for Walmart")  
qqline(residuals(model))



## Zehrs

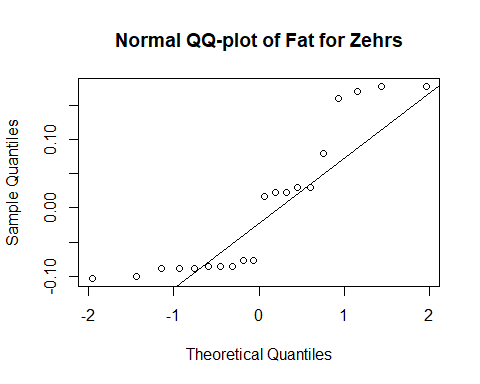
# Price\_per\_Z\_mL vs Calories

model <- lm(Price\_per\_Z\_mL ~ Calories\_per\_Z\_mL, data=subset(data, Store == 'Zehrs'))  
  
qqnorm(residuals(model), main="Normal QQ-plot of Calories for Zehrs")  
qqline(residuals(model))



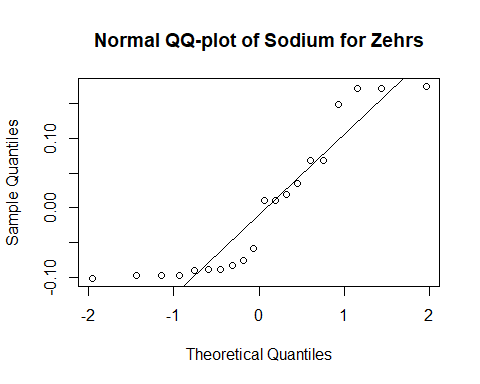
# Price\_per\_Z\_mL vs Fat (g)

model <- lm(Price\_per\_Z\_mL ~ Fat\_g\_per\_Z\_mL, data=subset(data, Store == 'Zehrs'))  
  
qqnorm(residuals(model), main="Normal QQ-plot of Fat for Zehrs")  
qqline(residuals(model))



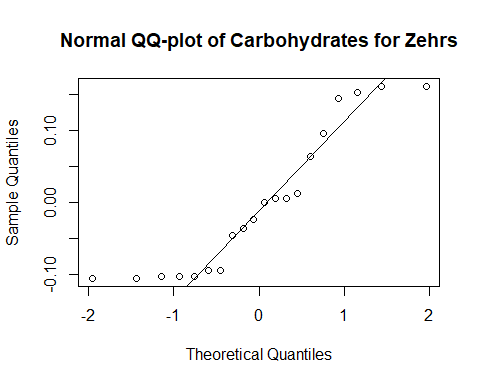
# Price\_per\_Z\_mL vs Sodium (mg)

model <- lm(Price\_per\_Z\_mL ~ Sodium\_mg\_per\_Z\_mL, data=subset(data, Store == 'Zehrs'))  
  
qqnorm(residuals(model), main="Normal QQ-plot of Sodium for Zehrs")  
qqline(residuals(model))



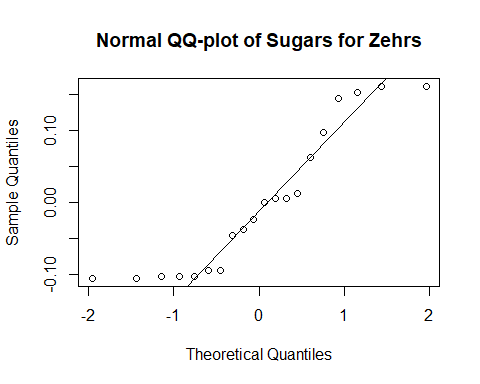
# Price\_per\_Z\_mL vs Carbohydrates (g)

model <- lm(Price\_per\_Z\_mL ~ Carbohydrates\_g\_per\_Z\_mL, data=subset(data, Store == 'Zehrs'))  
  
qqnorm(residuals(model), main="Normal QQ-plot of Carbohydrates for Zehrs")  
qqline(residuals(model))



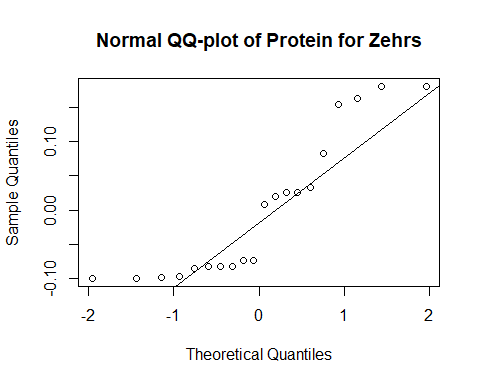
# Price\_per\_Z\_mL vs Sugars (g)

model <- lm(Price\_per\_Z\_mL ~ Sugars\_g\_per\_Z\_mL, data=subset(data, Store == 'Zehrs'))  
  
qqnorm(residuals(model), main="Normal QQ-plot of Sugars for Zehrs")  
qqline(residuals(model))



# Price\_per\_Z\_mL vs Protein (g)

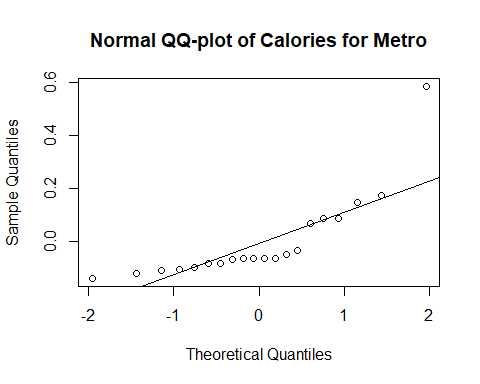
model <- lm(Price\_per\_Z\_mL ~ Protein\_g\_per\_Z\_mL, data=subset(data, Store == 'Zehrs'))  
  
qqnorm(residuals(model), main="Normal QQ-plot of Protein for Zehrs")  
qqline(residuals(model))



## Metro

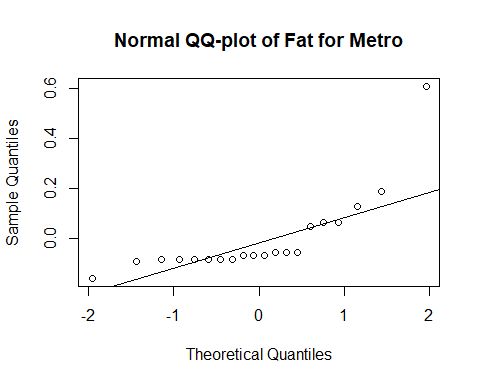
# Price\_per\_Z\_mL vs Calories

model <- lm(Price\_per\_Z\_mL ~ Calories\_per\_Z\_mL, data=subset(data, Store == 'Metro'))  
  
qqnorm(residuals(model), main="Normal QQ-plot of Calories for Metro")  
qqline(residuals(model))



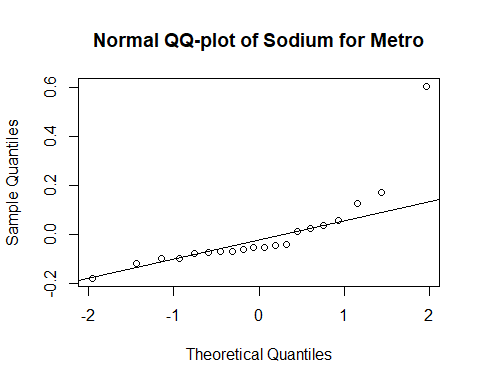
# Price\_per\_Z\_mL vs Fat (g)

model <- lm(Price\_per\_Z\_mL ~ Fat\_g\_per\_Z\_mL, data=subset(data, Store == 'Metro'))  
  
qqnorm(residuals(model), main="Normal QQ-plot of Fat for Metro")  
qqline(residuals(model))



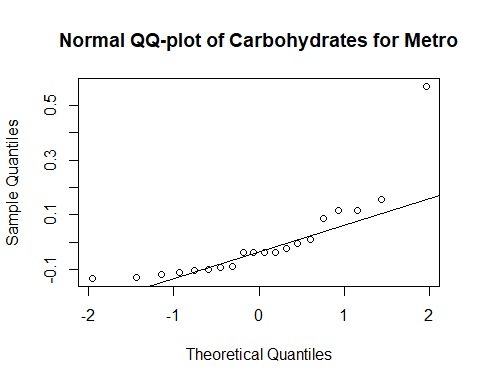
# Price\_per\_Z\_mL vs Sodium (mg)

model <- lm(Price\_per\_Z\_mL ~ Sodium\_mg\_per\_Z\_mL, data=subset(data, Store == 'Metro'))  
  
qqnorm(residuals(model), main="Normal QQ-plot of Sodium for Metro")  
qqline(residuals(model))



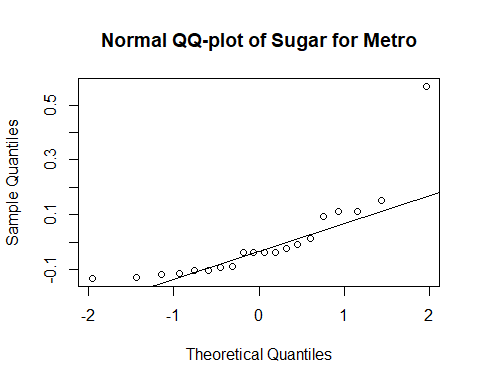
# Price\_per\_Z\_mL vs Carbohydrates (g)

model <- lm(Price\_per\_Z\_mL ~ Carbohydrates\_g\_per\_Z\_mL, data=subset(data, Store == 'Metro'))  
  
qqnorm(residuals(model), main="Normal QQ-plot of Carbohydrates for Metro")  
qqline(residuals(model))



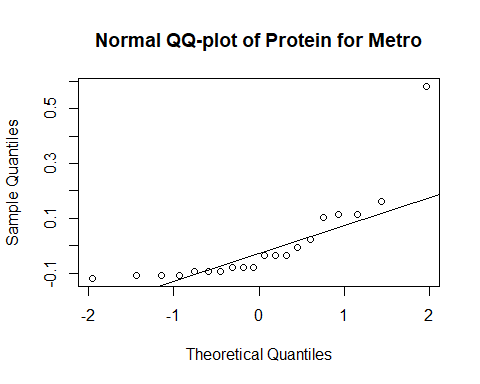
# Price\_per\_Z\_mL vs Sugars (g)

model <- lm(Price\_per\_Z\_mL ~ Sugars\_g\_per\_Z\_mL, data=subset(data, Store == 'Metro'))  
  
qqnorm(residuals(model), main="Normal QQ-plot of Sugar for Metro")  
qqline(residuals(model))



# Price\_per\_Z\_mL vs Protein (g)

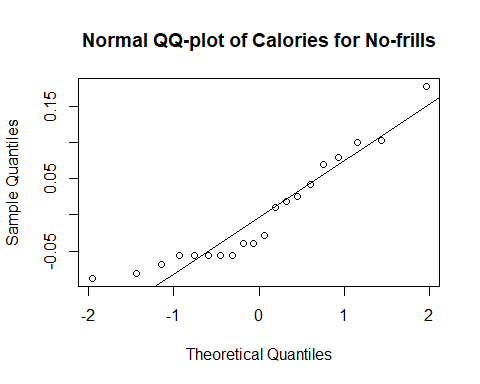
model <- lm(Price\_per\_Z\_mL ~ Protein\_g\_per\_Z\_mL, data=subset(data, Store == 'Metro'))  
  
qqnorm(residuals(model), main="Normal QQ-plot of Protein for Metro")  
qqline(residuals(model))



## No-frills

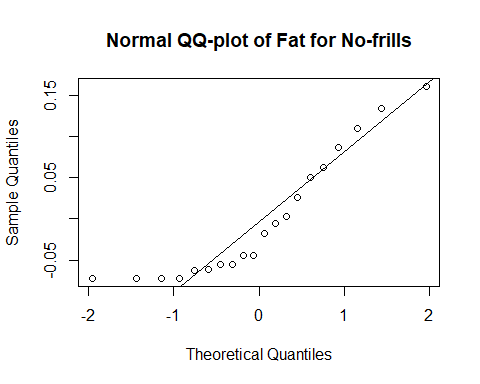
# Price\_per\_Z\_mL vs Calories

model <- lm(Price\_per\_Z\_mL ~ Calories\_per\_Z\_mL, data=subset(data, Store == 'No-Frills'))  
  
qqnorm(residuals(model), main="Normal QQ-plot of Calories for No-frills")  
qqline(residuals(model))



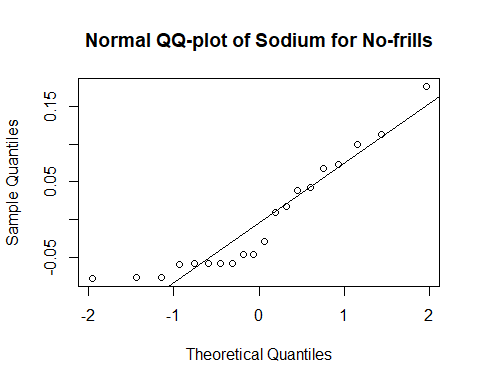
# Price\_per\_Z\_mL vs Fat (g)

model <- lm(Price\_per\_Z\_mL ~ Fat\_g\_per\_Z\_mL, data=subset(data, Store == 'No-Frills'))  
  
qqnorm(residuals(model), main="Normal QQ-plot of Fat for No-frills")  
qqline(residuals(model))



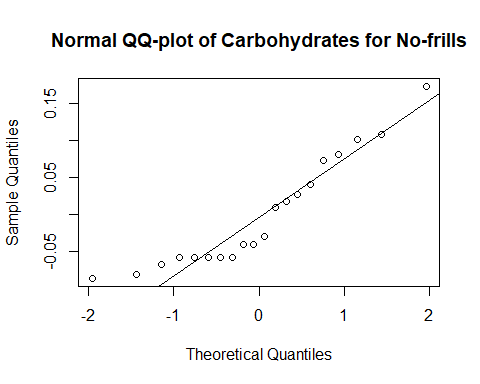
# Price\_per\_Z\_mL vs Sodium (mg)

model <- lm(Price\_per\_Z\_mL ~ Sodium\_mg\_per\_Z\_mL, data=subset(data, Store == 'No-Frills'))  
  
qqnorm(residuals(model), main="Normal QQ-plot of Sodium for No-frills")  
qqline(residuals(model))



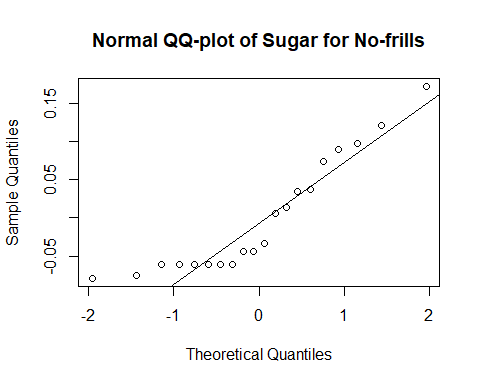
# Price\_per\_Z\_mL vs Carbohydrates (g)

model <- lm(Price\_per\_Z\_mL ~ Carbohydrates\_g\_per\_Z\_mL, data=subset(data, Store == 'No-Frills'))  
  
qqnorm(residuals(model), main="Normal QQ-plot of Carbohydrates for No-frills")  
qqline(residuals(model))



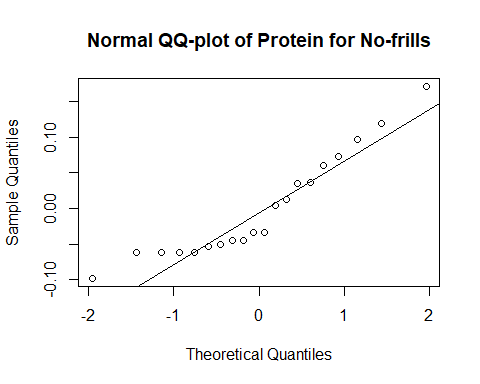
# Price\_per\_Z\_mL vs Sugars (g)

model <- lm(Price\_per\_Z\_mL ~ Sugars\_g\_per\_Z\_mL, data=subset(data, Store == 'No-Frills'))  
  
qqnorm(residuals(model), main="Normal QQ-plot of Sugar for No-frills")  
qqline(residuals(model))



# Price\_per\_Z\_mL vs Protein (g)

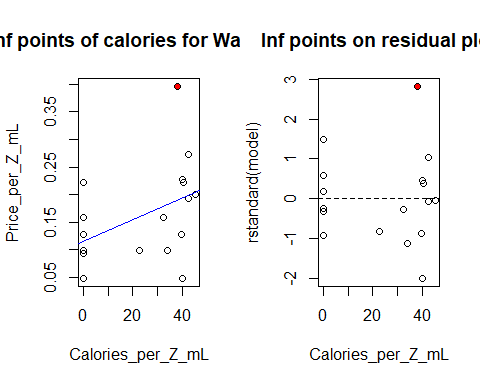
model <- lm(Price\_per\_Z\_mL ~ Protein\_g\_per\_Z\_mL, data=subset(data, Store == 'No-Frills'))  
  
qqnorm(residuals(model), main="Normal QQ-plot of Protein for No-frills")  
qqline(residuals(model))



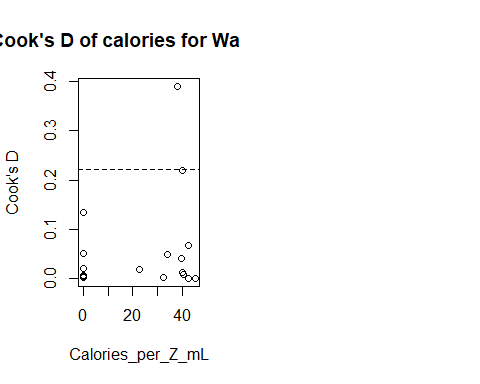
### Influential points

We are going to find the influential points for the 3 models we have selected

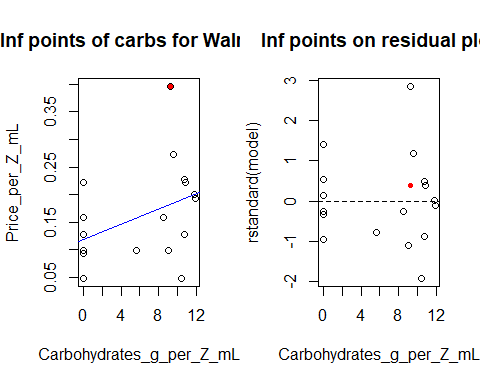
model <- lm(Price\_per\_Z\_mL ~ Calories\_per\_Z\_mL, data=subset(data, Store == 'Walmart'))  
  
D <- cooks.distance(model)  
cutoff <- 4/nrow(subset(data, Store == 'Walmart'))  
  
ynew <- subset(data, Store == 'Walmart')$Price\_per\_Z\_mL  
  
ri <- rstandard(model)  
  
par(mfrow=c(1,2))  
  
plot(Price\_per\_Z\_mL ~ Calories\_per\_Z\_mL, main="Inf points of calories for Walmart", data=subset(data, Store == 'Walmart'))  
  
abline(a = model$coefficients[1], b = model$coefficients[2], col="blue")  
points(subset(data, Store == 'Walmart')$Calories\_per\_Z\_mL[which(D > 4/(nrow(subset(data, Store == 'Walmart'))-2))], ynew[which(D > 4/(nrow(subset(data, Store == 'Walmart'))-2))], col="red", pch=20)  
  
  
  
plot(rstandard(model) ~ Calories\_per\_Z\_mL, data=subset(data, Store == 'Walmart'), main="Inf points on residual plot")  
  
points(subset(data, Store == 'Walmart')$Calories\_per\_Z\_mL[which(D > 4/(nrow(subset(data, Store == 'Walmart'))-2))], rstandard(model)[which(D > 4/(nrow(subset(data, Store == 'Walmart'))-2))], col="red", pch=20)  
  
abline(h = 0, lty = 2)



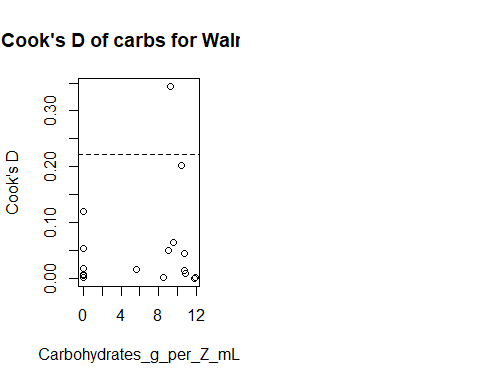
plot(D ~ Calories\_per\_Z\_mL, ylab="Cook's D", main = "Cook's D of calories for Walmart", data=subset(data, Store == 'Walmart'))  
  
abline(h = (4/(nrow(subset(data, Store == 'Walmart'))-2)), lty=2)



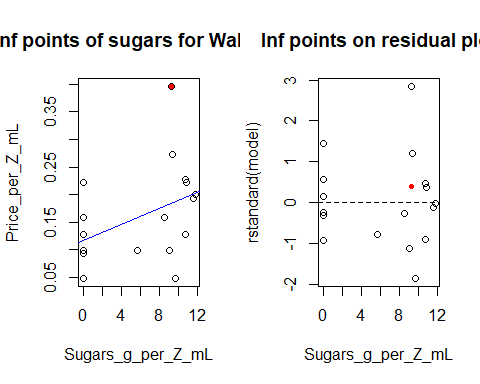
model <- lm(Price\_per\_Z\_mL ~ Carbohydrates\_g\_per\_Z\_mL, data=subset(data, Store == 'Walmart'))  
  
D <- cooks.distance(model)  
cutoff <- 4/nrow(subset(data, Store == 'Walmart'))  
  
ynew <- subset(data, Store == 'Walmart')$Price\_per\_Z\_mL  
  
ri <- rstandard(model)  
  
par(mfrow=c(1,2))  
  
plot(Price\_per\_Z\_mL ~ Carbohydrates\_g\_per\_Z\_mL, main="Inf points of carbs for Walmart", data=subset(data, Store == 'Walmart'))  
  
abline(a = model$coefficients[1], b = model$coefficients[2], col="blue")  
  
points(subset(data, Store == 'Walmart')$Carbohydrates\_g\_per\_Z\_mL[which(D > 4/(nrow(subset(data, Store == 'Walmart'))-2))], ynew[which(D > 4/(nrow(subset(data, Store == 'Walmart'))-2))], col="red", pch=20)  
  
  
  
plot(rstandard(model) ~ Carbohydrates\_g\_per\_Z\_mL, data=subset(data, Store == 'Walmart'), main="Inf points on residual plot")  
  
points(subset(data, Store == 'Walmart')$Carbohydrates\_g\_per\_Z\_mL[which(D > 4/(nrow(subset(data, Store == 'Walmart'))-2))], ynew[which(D > 4/(nrow(subset(data, Store == 'Walmart'))-2))], col="red", pch=20)  
  
abline(h = 0, lty = 2)



plot(D ~ Carbohydrates\_g\_per\_Z\_mL, ylab="Cook's D", main = "Cook's D of carbs for Walmart", data=subset(data, Store == 'Walmart'))  
  
abline(h = (4/(nrow(subset(data, Store == 'Walmart'))-2)), lty=2)



model <- lm(Price\_per\_Z\_mL ~ Sugars\_g\_per\_Z\_mL, data=subset(data, Store == 'Walmart'))  
  
D <- cooks.distance(model)  
cutoff <- 4/nrow(subset(data, Store == 'Walmart'))  
  
ynew <- subset(data, Store == 'Walmart')$Price\_per\_Z\_mL  
  
ri <- rstandard(model)  
  
par(mfrow=c(1,2))  
  
plot(Price\_per\_Z\_mL ~ Sugars\_g\_per\_Z\_mL, main="Inf points of sugars for Walmart", data=subset(data, Store == 'Walmart'))  
  
abline(a = model$coefficients[1], b = model$coefficients[2], col="blue")  
points(subset(data, Store == 'Walmart')$Sugars\_g\_per\_Z\_mL[which(D > 4/(nrow(subset(data, Store == 'Walmart'))-2))], ynew[which(D > 4/(nrow(subset(data, Store == 'Walmart'))-2))], col="red", pch=20)  
  
  
  
plot(rstandard(model) ~ Sugars\_g\_per\_Z\_mL, data=subset(data, Store == 'Walmart'), main="Inf points on residual plot")  
  
points(subset(data, Store == 'Walmart')$Sugars\_g\_per\_Z\_mL[which(D > 4/(nrow(subset(data, Store == 'Walmart'))-2))], ynew[which(D > 4/(nrow(subset(data, Store == 'Walmart'))-2))], col="red", pch=20)  
  
abline(h = 0, lty = 2)



plot(D ~ Sugars\_g\_per\_Z\_mL, ylab="Cook's D", main = "Cook's D of sugars for Walmart", data=subset(data, Store == 'Walmart'))  
  
abline(h = (4/(nrow(subset(data, Store == 'Walmart'))-2)), lty=2)

