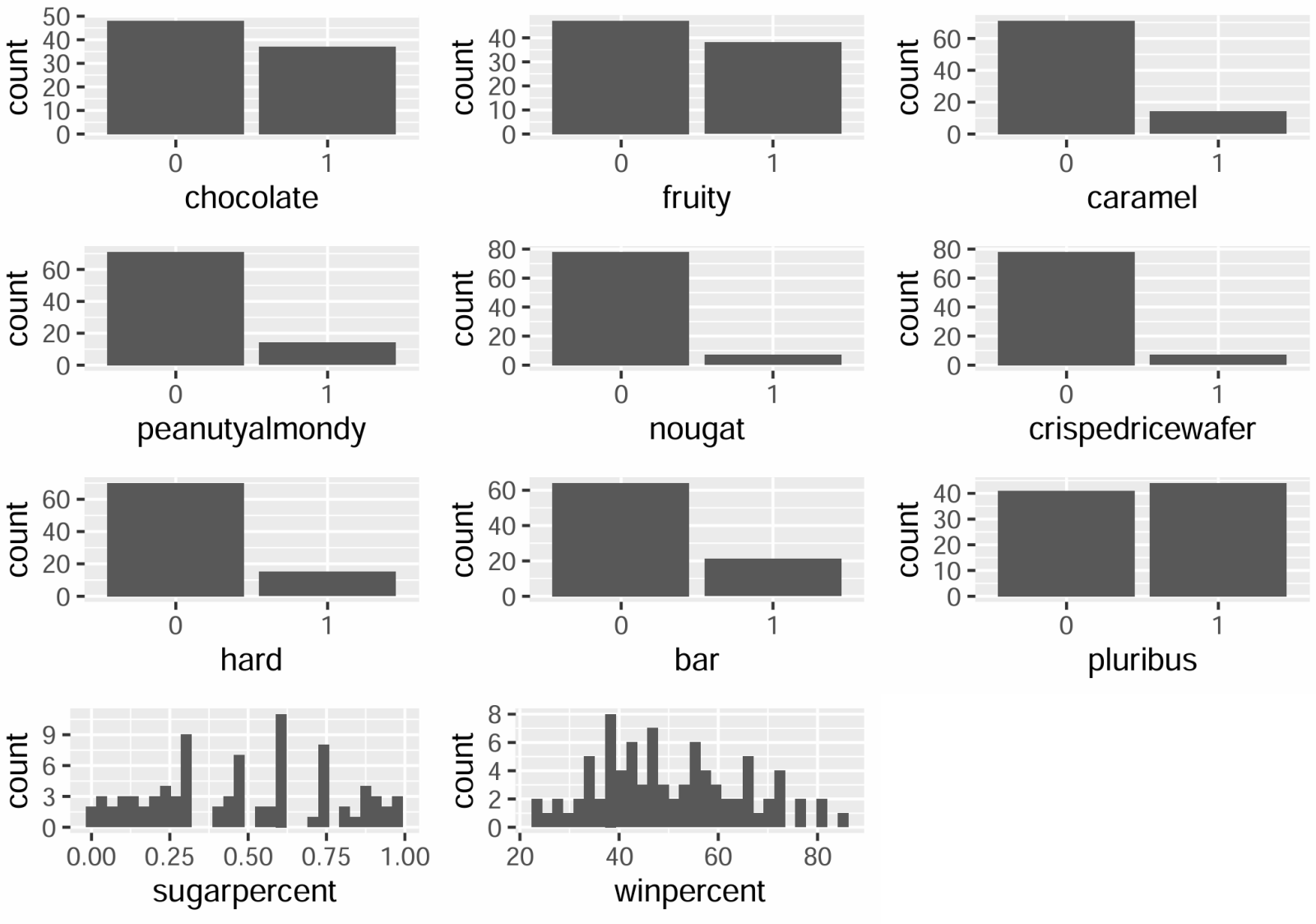
What makes a Halloween candy desirable?

ANAV VORA AND LAVANYA KUDLI



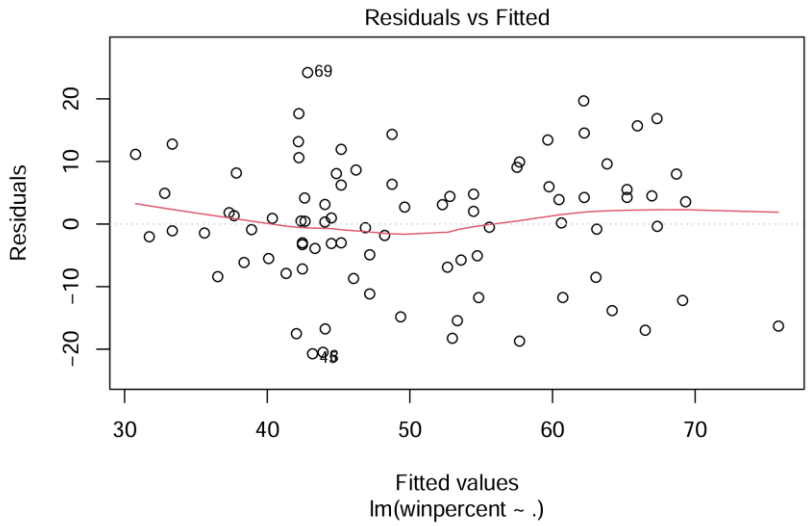
Potential Influencers

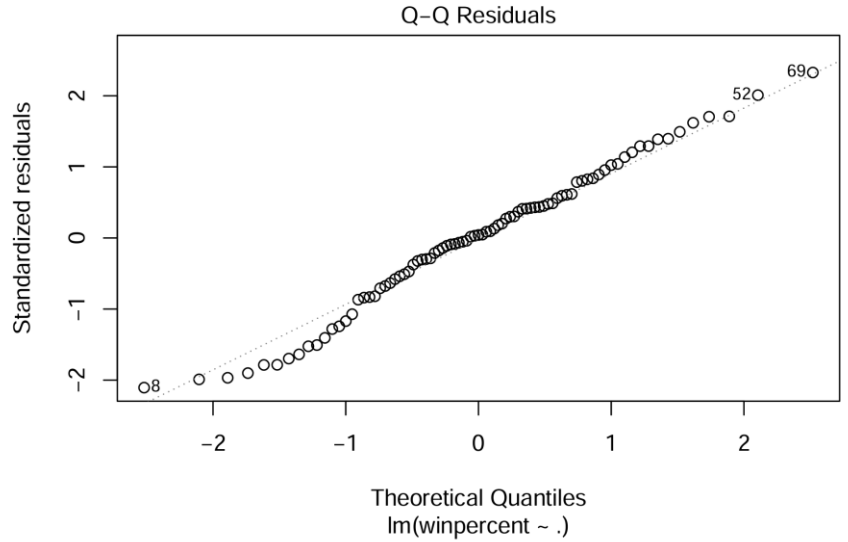
Model Fitting

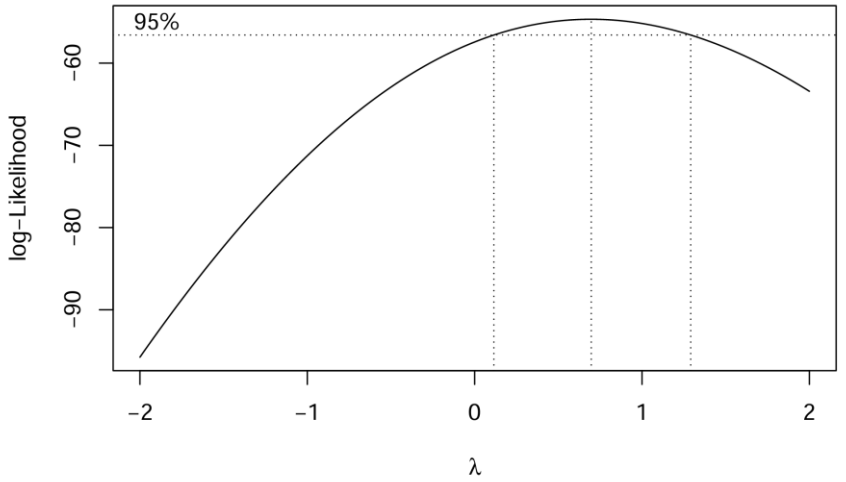
The overall F-test of the model reveals that at least one of the betas is not zero (p-value of overall F-test = 5.25 x 10^{-9}

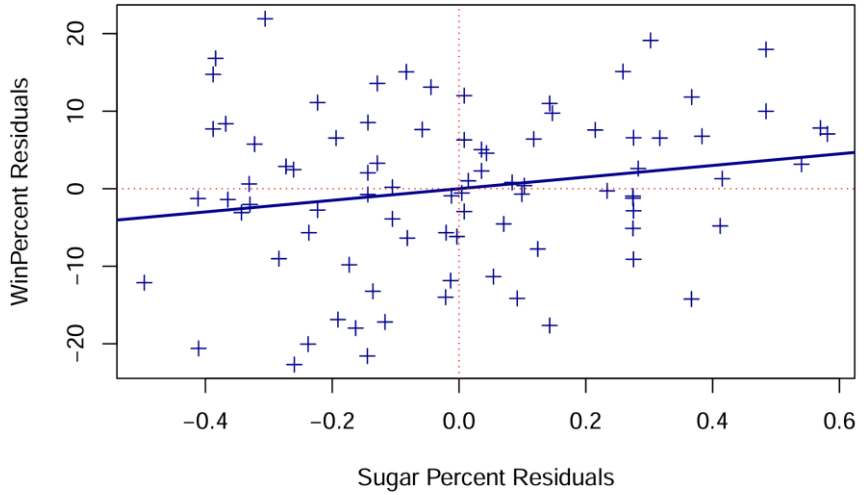
```
Call:
lm(formula = winpercent ~ ., data = Candy_Data_Subset)

Residuals:
     Min       1Q   Median       3Q      Max
-20.7320  -6.1438   0.4359   6.2276  24.2048

Coefficients:
                  Estimate Std. Error t value Pr(>|t|)
(Intercept)         33.262      4.159   7.997 1.31e-11 ***
chocolate1          19.216      3.871   4.964 4.30e-06 ***
fruity1              9.587      3.764   2.547  0.01294 *
caramel1             1.982      3.654   0.542  0.58918
peanutyalmondy1      9.454      3.574   2.645  0.00996 **
nougat1              1.918      5.628   0.341  0.73422
crispedricewafer1    8.637      5.267   1.640  0.10529
hard1               -5.866      3.448  -1.701  0.09306 .
bar1                -1.233      4.821  -0.256  0.79879
pluribus1           -1.147      3.031  -0.378  0.70629
sugarpercent         7.490      4.421   1.694  0.09444 .
```

Diagnostics









Fine Tuning and Interpretation

```r
#Now dropping peanutyalmondy also (next largest p value)
candy.red9 = lm(winpercent ~ chocolate, data=Candy_Data_Subset)
anova(candy.red9,Candy.mlr)
```

```
## Analysis of Variance Table
##
## Model 1: winpercent ~ chocolate
## Model 2: winpercent ~ chocolate + fruity + caramel + peanutyalmondy +
##     nougat + crispedricewafer + hard + bar + pluribus + sugarpercent
##   Res.Df     RSS Df Sum of Sq      F  Pr(>F)
## 1     83 10818.5
## 2     74  8494.7  9    2323.8 2.2493 0.02774 *
## ---
## Signif. codes:  0 ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’.’ 0.1 ’ ’ 1
```

```r
summary(candy.final)
```

```
##
## Call:
## lm(formula = winpercent ~ chocolate + peanutyalmondy, data = Candy_Data_Subset)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -26.0296  -7.6657   0.0797   7.3629  25.2130
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)       41.825      1.619  25.828  < 2e-16 ***
## chocolate1        16.625      2.640   6.297 1.42e-08 ***
## peanutyalmondy1    7.623      3.529   2.160   0.0337 *
## ---
## Signif. codes:  0 ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’.’ 0.1 ’ ’ 1
##
## Residual standard error: 11.17 on 82 degrees of freedom
## Multiple R-squared:  0.4372, Adjusted R-squared:  0.4235
## F-statistic: 31.85 on 2 and 82 DF,  p-value: 5.822e-11
```