101A - HW3

```
library(car)
```

```
## Loading required package: carData
```

```
## Warning: package 'carData' was built under R version 4.1.2
```

```
library(plyr)
```

```
birth <- read.csv("Births5000.csv")
bmean <- aggregate(birth$Birth.weight.g, list(birth$Mother.Minority),mean)
bmean
```

```
##     Group.1        x
## 1 Nonwhite 5.767484
## 2    White 6.175173
```

```
bvar <- aggregate(birth$Birth.weight.g, list(birth$Mother.Minority),var)
bvar
```

```
##     Group.1        x
## 1 Nonwhite 1.784069
## 2    White 1.424155
```

```
model <- lm(birth$Birth.Weight..g.~birth$Mother.Minority)
summary(model)
```

```
##
## Call:
## lm(formula = birth$Birth.Weight..g. ~ birth$Mother.Minority)
##
## Residuals:
##       Min       1Q   Median       3Q      Max
## -3060.52  -308.14    32.36   372.86  1990.23
##
## Coefficients:
##                             Estimate Std. Error t value Pr(>|t|)
## (Intercept)                  3135.97      18.02  173.99   <2e-16 ***
## birth$Mother.MinorityWhite    208.30      20.41   10.21   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 598.1 on 4998 degrees of freedom
## Multiple R-squared:  0.02041,    Adjusted R-squared:  0.02022
## F-statistic: 104.2 on 1 and 4998 DF,  p-value: < 2.2e-16
```

```
#SLR - Slope = 208.3 Intercept = 3135.97 Standard Error = 598.1 t-statistic - 10.21 d.f = 4998 p-value

ttest<- t.test(birth$Birth.Weight..g. ~ birth$Mother.Minority, alternative = "two.sided", var.equal = T
ttest
```

```
##
##  Two Sample t-test
##
## data:  birth$Birth.Weight..g. by birth$Mother.Minority
## t = -10.206, df = 4998, p-value < 2.2e-16
## alternative hypothesis: true difference in means between group Nonwhite and group White is not equal
## 95 percent confidence interval:
##  -248.3168 -168.2899
## sample estimates:
## mean in group Nonwhite    mean in group White
##                3135.966               3344.269
```

```
#Two sided t-test -
```
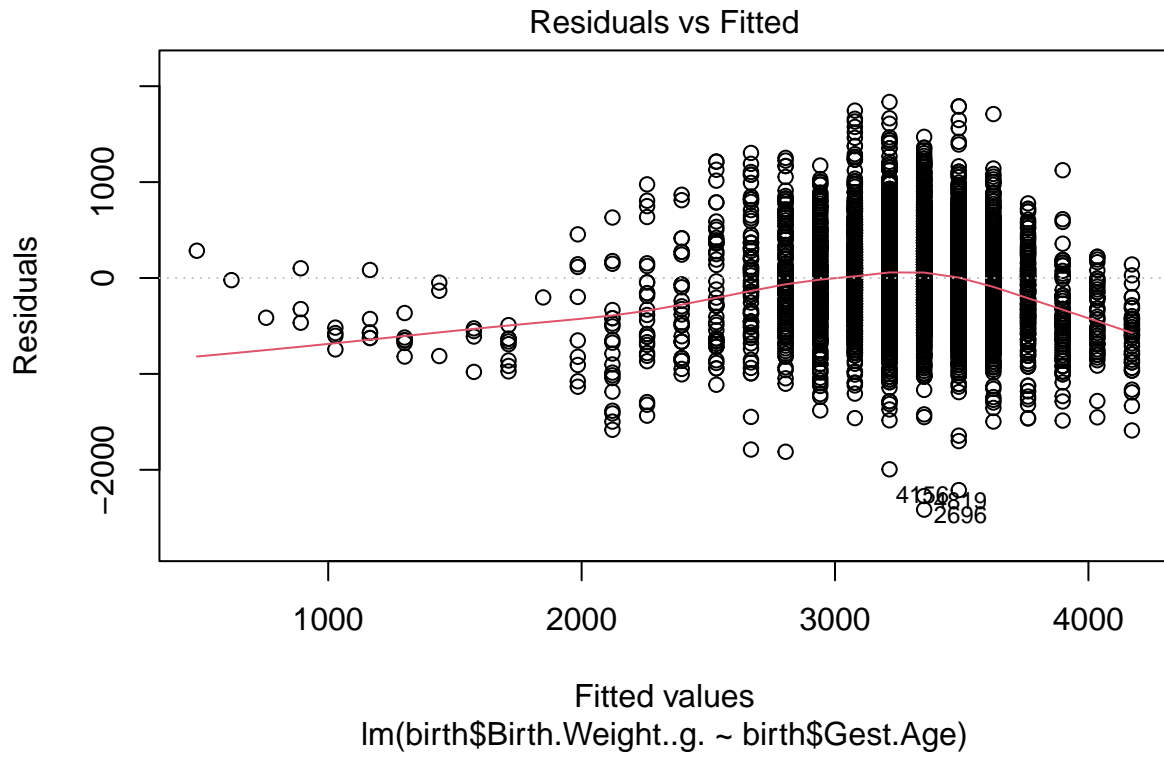
```
plot(model)
```


Residuals vs Fitted

Fitted values
lm(birth$Birth.Weight..g. ~ birth$Mother.Minority)

2

Normal Q–Q

Standardized residuals

Theoretical Quantiles
lm(birth$Birth.Weight..g. ~ birth$Mother.Minority)

Scale−Location

Fitted values
lm(birth$Birth.Weight..g. ~ birth$Mother.Minority)

Residuals vs Leverage

Leverage
lm(birth$Birth.Weight..g. ~ birth$Mother.Minority)
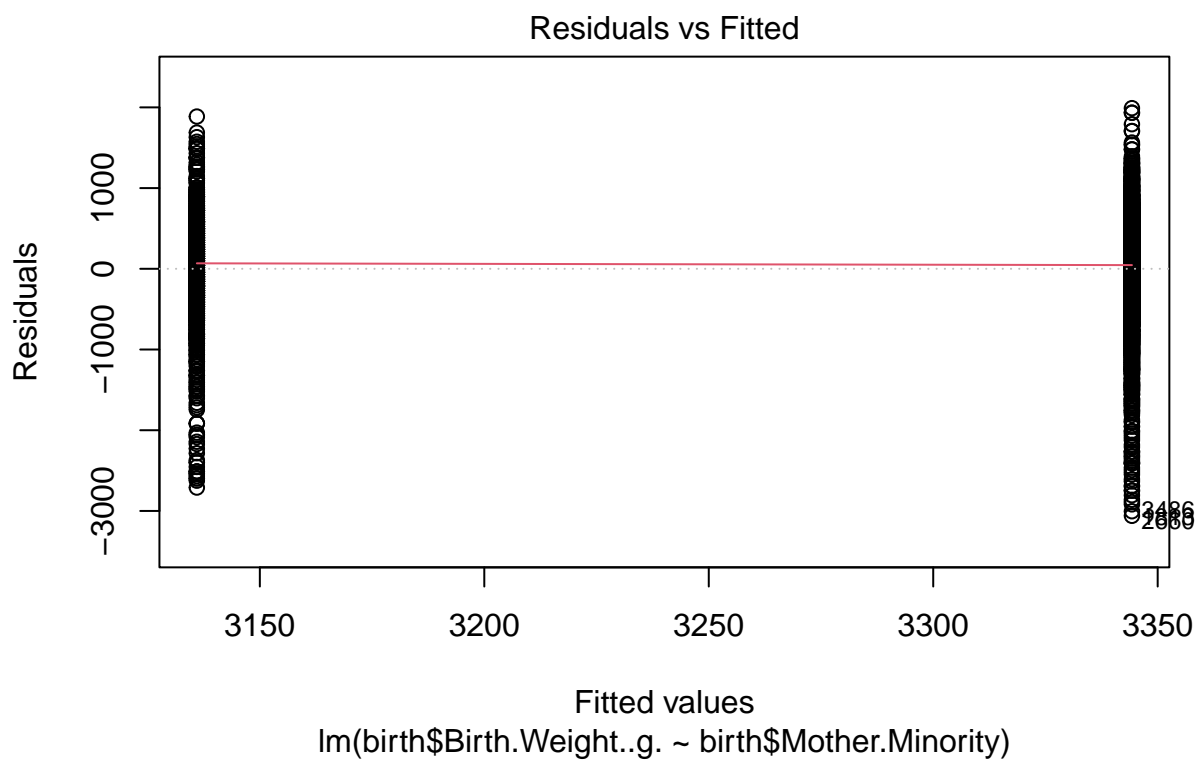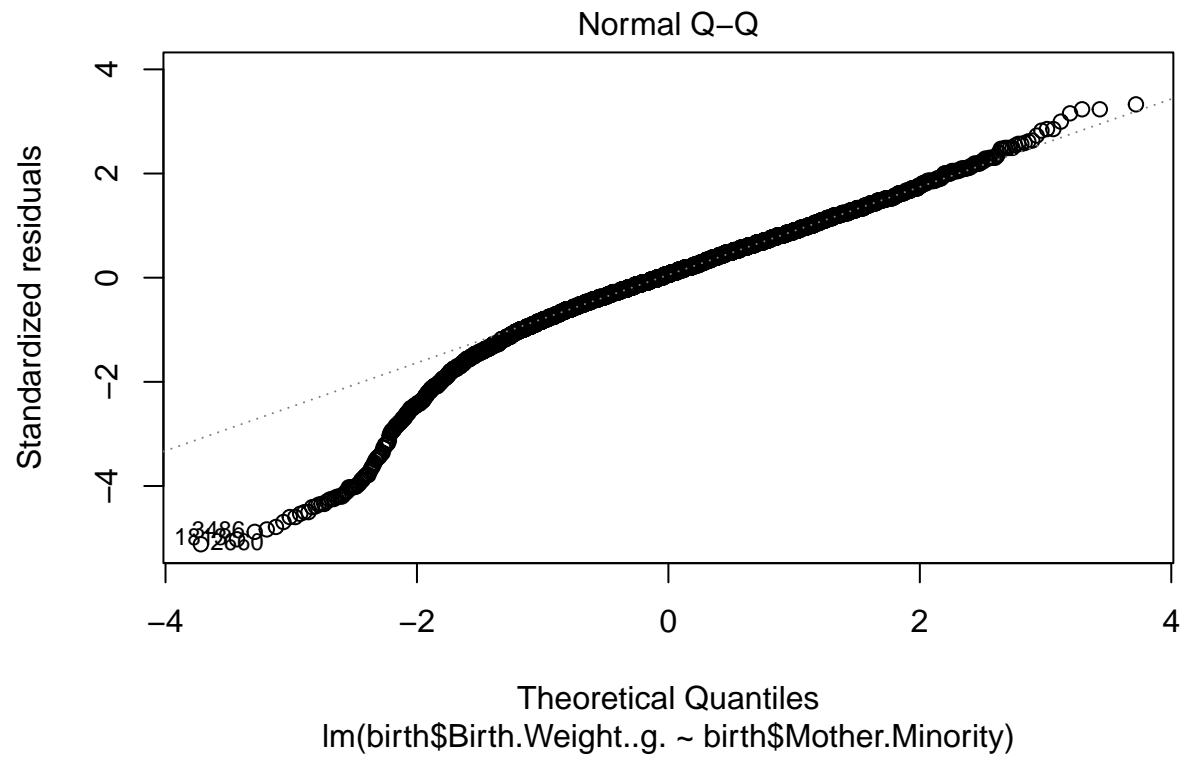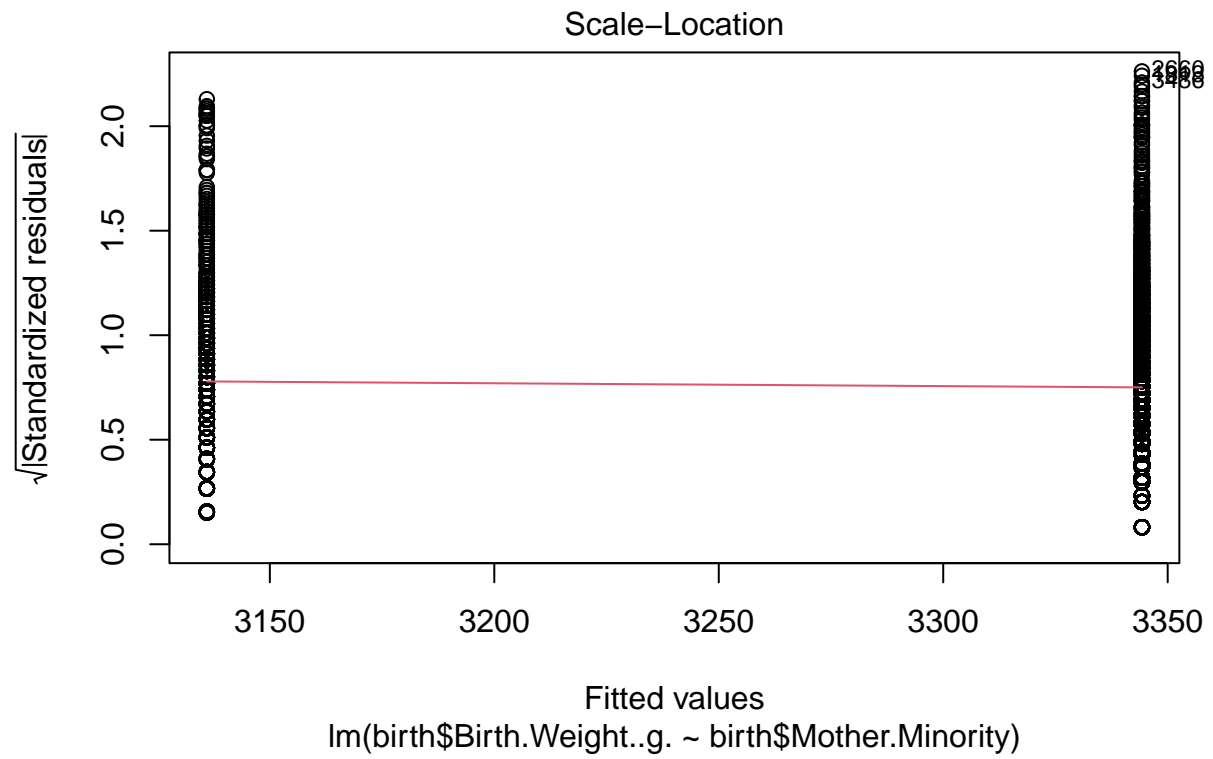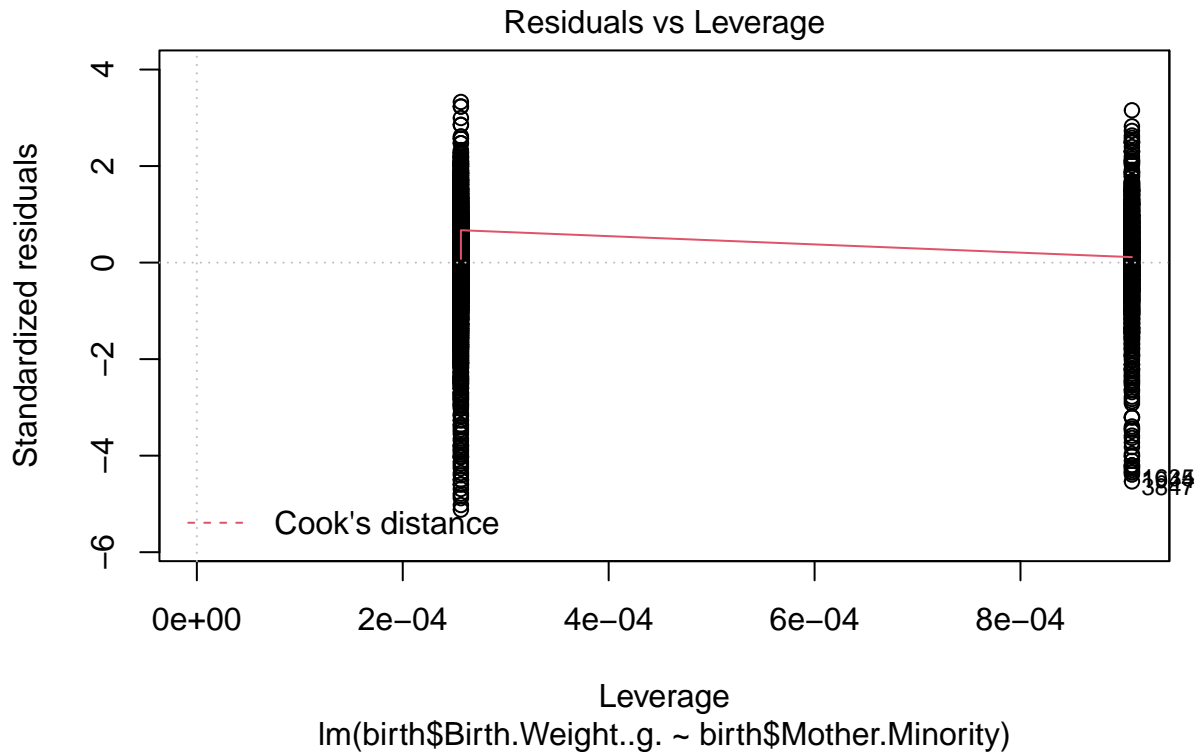
```
glm <- lm(birth$Birth.Weight..g. ~ birth$Gest.Age)
summary(glm)
```
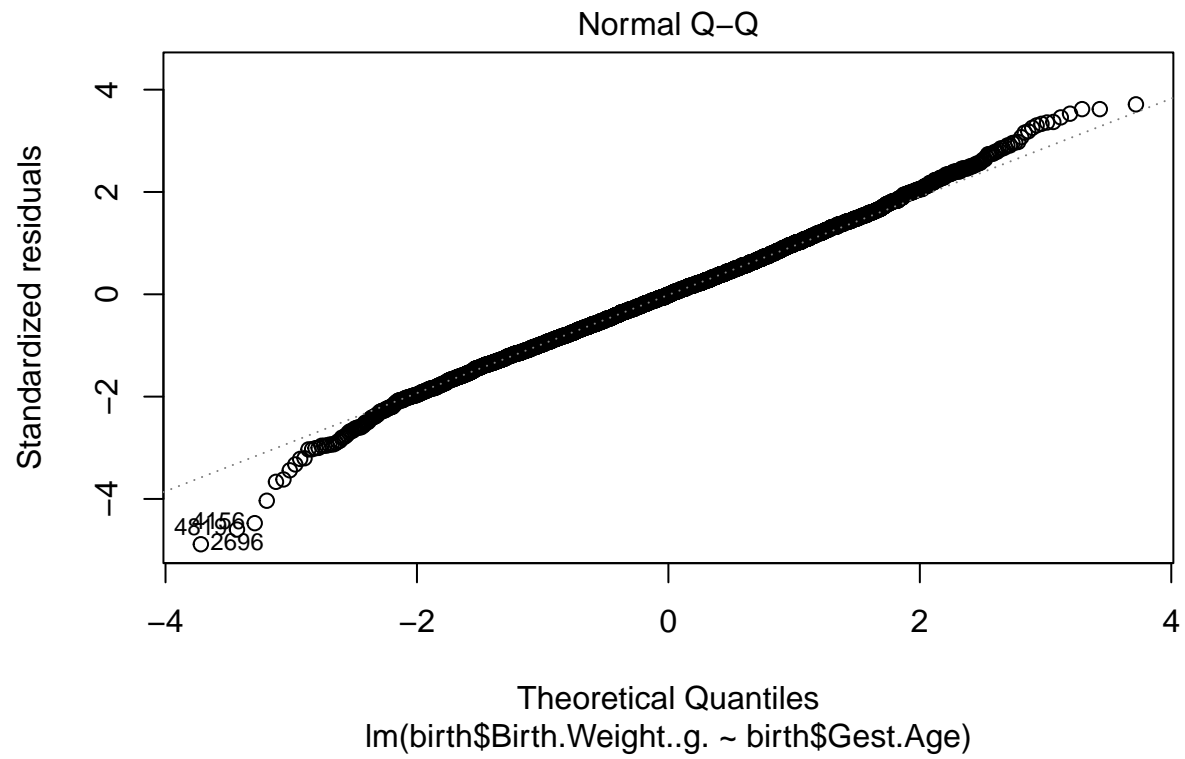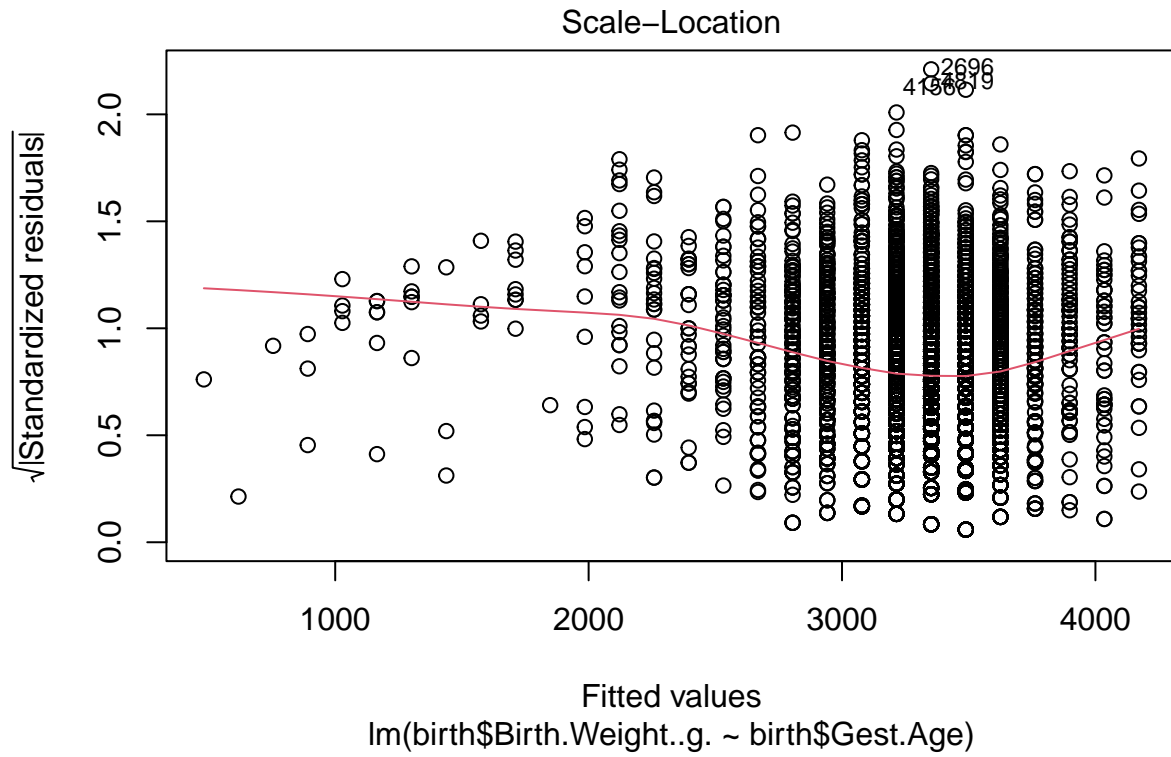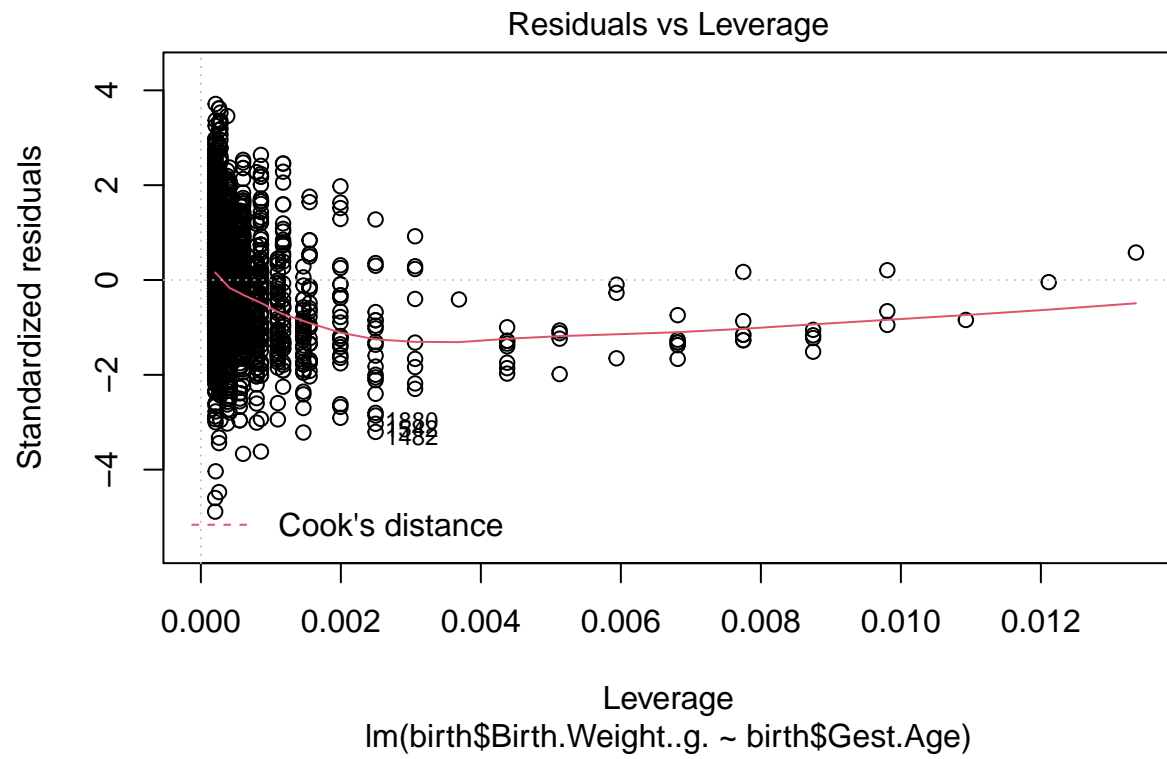
```
##
## Call:
## lm(formula = birth$Birth.Weight..g. ~ birth$Gest.Age)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2415.36  -326.01    -3.48   313.85  1835.69
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)     -1978.486    106.481  -18.58   <2e-16 ***
## birth$Gest.Age    136.672      2.752   49.66   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 494.4 on 4998 degrees of freedom
## Multiple R-squared:  0.3304, Adjusted R-squared:  0.3303
## F-statistic:  2467 on 1 and 4998 DF,  p-value: < 2.2e-16
```

```
plot(glm)
```

Residuals vs Fitted

Residuals

Fitted values
lm(birth$Birth.Weight..g. ~ birth$Gest.Age)

Normal Q–Q

Theoretical Quantiles
lm(birth$Birth.Weight..g. ~ birth$Gest.Age)

Scale−Location

√|Standardized residuals|

Fitted values
lm(birth$Birth.Weight..g. ~ birth$Gest.Age)

Residuals vs Leverage

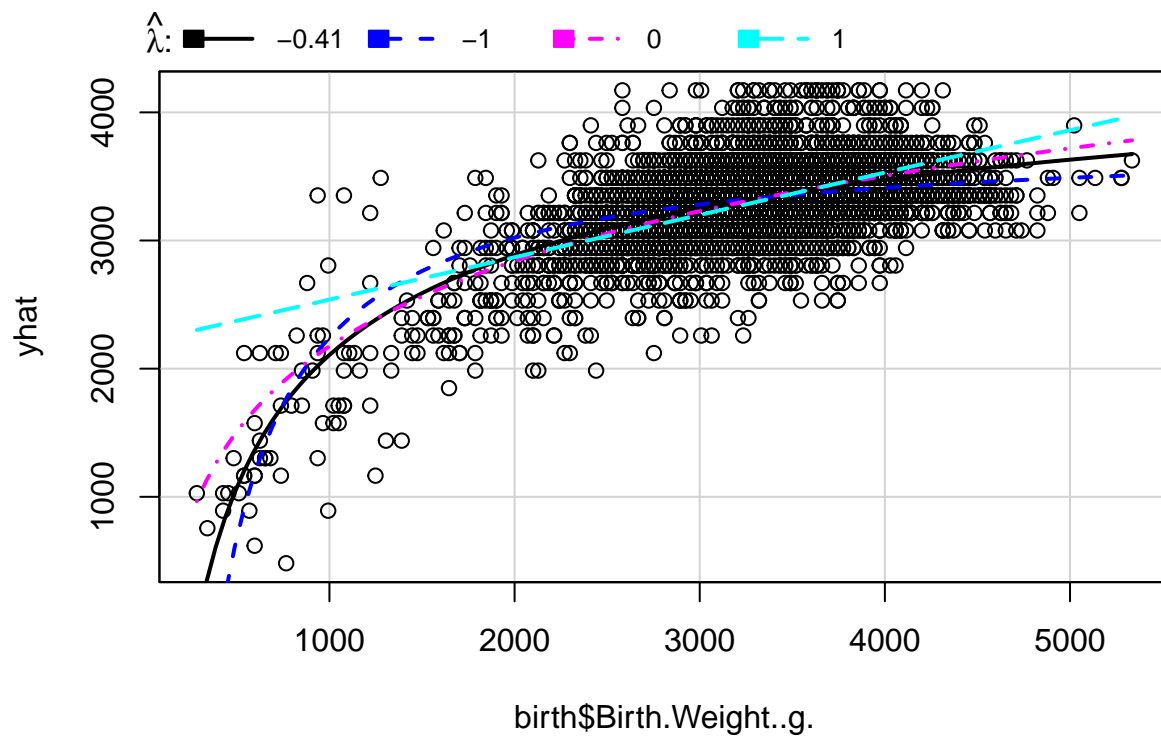lm(birth$Birth.Weight..g. ~ birth$Gest.Age)

```
invglm <- invResPlot(glm)
```

```
summary(invglm)
```

```
##      lambda              RSS
##  Min.   :-1.0000   Min.   :337589479
##  1st Qu.:-0.5540   1st Qu.:344376125
##  Median :-0.2027   Median :353915275
##  Mean   :-0.1013   Mean   :362291697
##  3rd Qu.: 0.2500   3rd Qu.:371830847
##  Max.   : 1.0000   Max.   :403746760
```

```
iglm <- lm(invglm)
```

```
test <- read.csv("CBDtestNoY.csv")
train <- read.csv("CBDtrain.csv")

dim(test)
```

```
## [1] 1155    21
```

```
dim(train)
```

```
## [1] 2000    22
```