kids

Chris Parrish

January 17, 2016

kids

Import the data.

```
data <- read.csv("Kids198.csv", header=TRUE)
head(data)
```

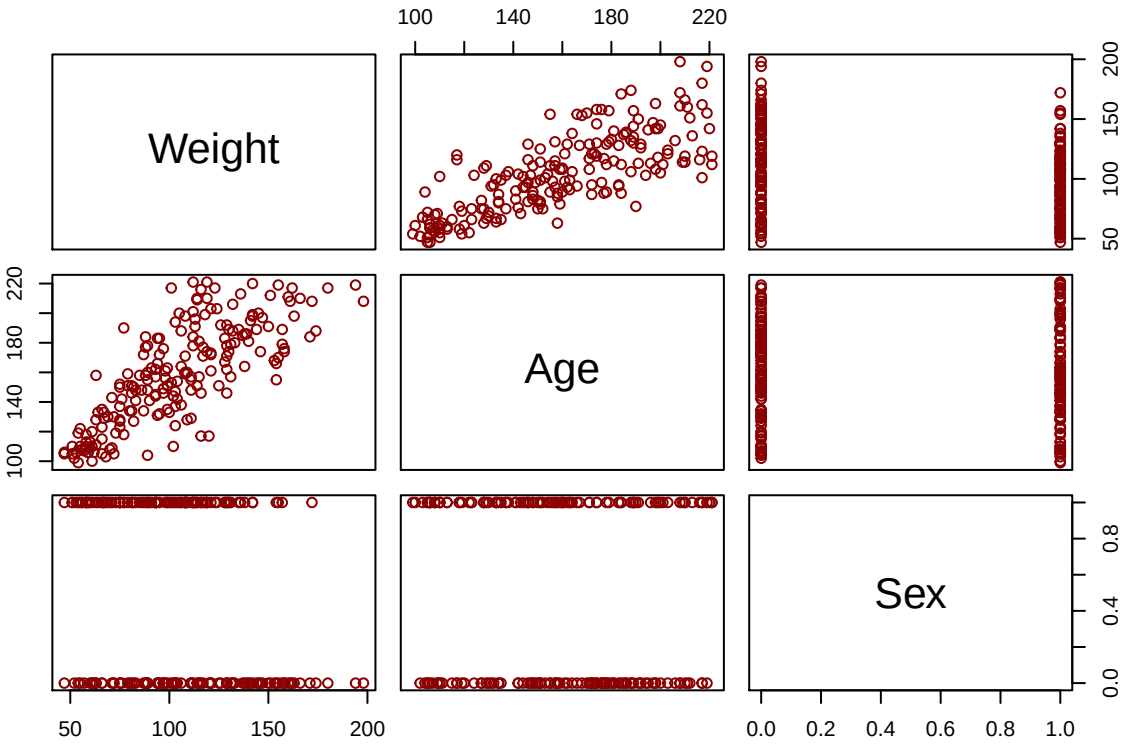
```
##   Height Weight Age Sex Race
## 1   67.8    166 210   0    1
## 2   63.0     93 144   1    0
## 3   50.1     54 119   0    0
## 4   55.7     69 130   1    0
## 5   63.2    115 157   0    0
## 6   48.8     52 102   0    0
```

```
dim(data)
```

```
## [1] 198   5
```

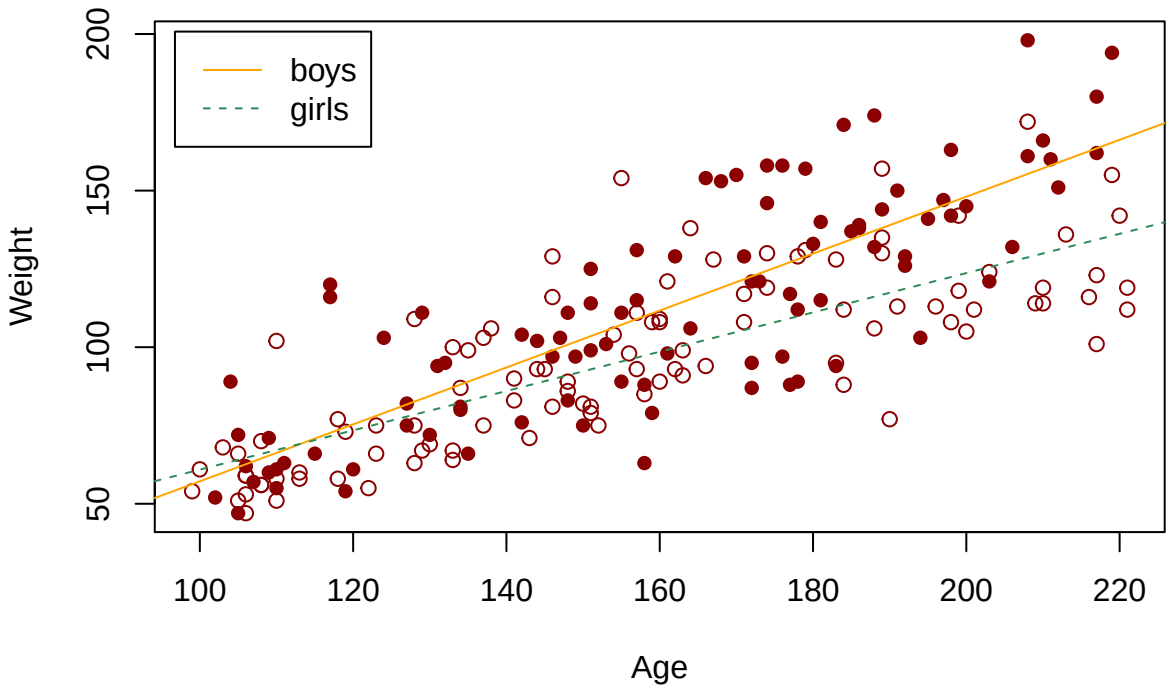
Scatterplot matrix.

```
pairs(~ Weight + Age + Sex, data=data, col="darkred")
```



Separate linear models for boys and girls

```
plot(Weight ~ Age, data=data,
     pch=16 - 15 * Sex, col="darkred")
boys.lm <- lm(Weight[Sex==0] ~ Age[Sex==0], data=data)
abline(boys.lm, col="orange", lty=1)      # boys
girls.lm <- lm(Weight[Sex==1] ~ Age[Sex==1], data=data)
abline(girls.lm, col="seagreen4", lty=2)  # girls
legend(x="topleft", legend=c("boys", "girls"),
       lty=1:2, col=c("orange", "seagreen4"), inset=0.02)
```



$\widehat{Weight}$ = -33.693 + 0.909 Age

```
coef(boys.lm)
```

```
##   (Intercept) Age[Sex == 0]
##   -33.6925380     0.9087098
```

$\widehat{Weight}$ = -1.842 + 0.627 Age

```
coef(girls.lm)
```

```
##   (Intercept) Age[Sex == 1]
##    -1.8419725     0.6274878
```

Multiple regression with interaction.

```
kids.lm <- lm(Weight ~ Age * Sex, data=data)
options(show.signif.stars=FALSE)
summary(kids.lm)
```

```
##
## Call:
## lm(formula = Weight ~ Age * Sex, data = data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -46.884 -12.055  -2.782  10.185  58.581
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -33.69254   10.00727  -3.367 0.000917
## Age           0.90871    0.06106  14.882  < 2e-16
## Sex          31.85057   13.24269   2.405 0.017106
## Age:Sex      -0.28122    0.08164  -3.445 0.000700
##
## Residual standard error: 19.19 on 194 degrees of freedom
## Multiple R-squared:  0.6683, Adjusted R-squared:  0.6631
## F-statistic: 130.3 on 3 and 194 DF,  p-value: < 2.2e-16
```

For boys:

$\widehat{Weight}$ = -33.693 + 0.909 Age

For girls:

$\widehat{Weight}$ = -1.842 + 0.627 Age

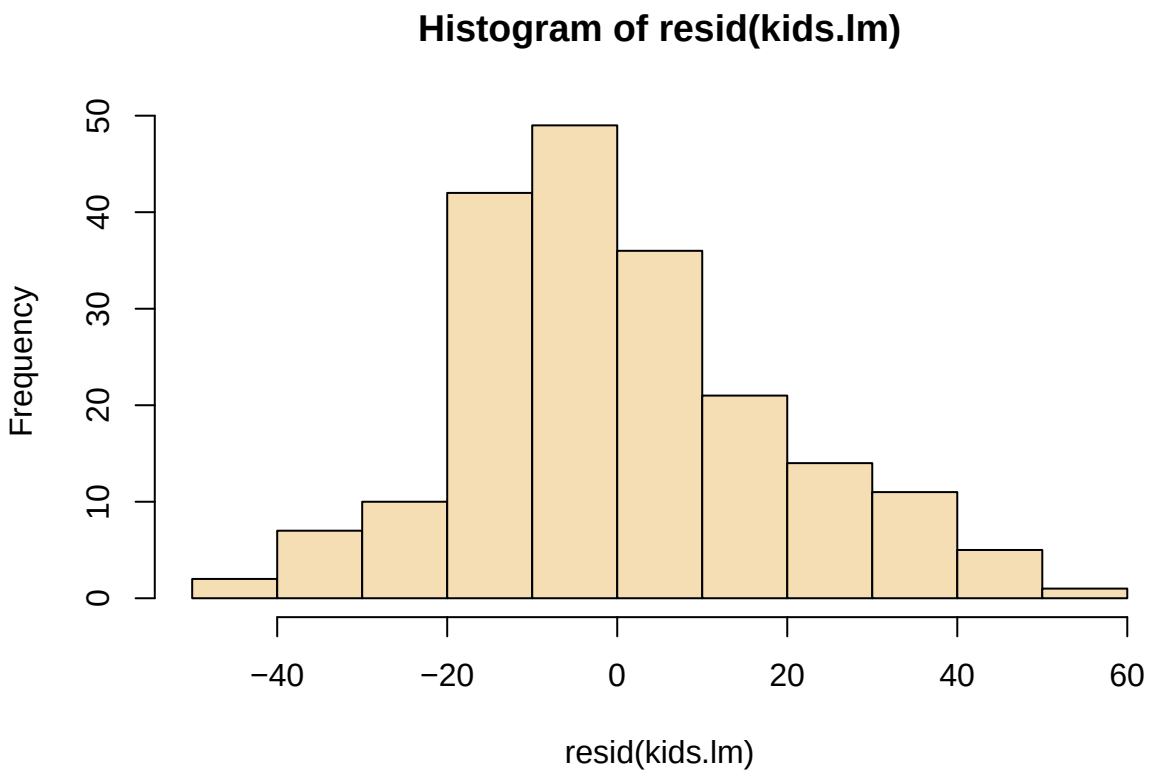
ANOVA.

```
anova(kids.lm)
```

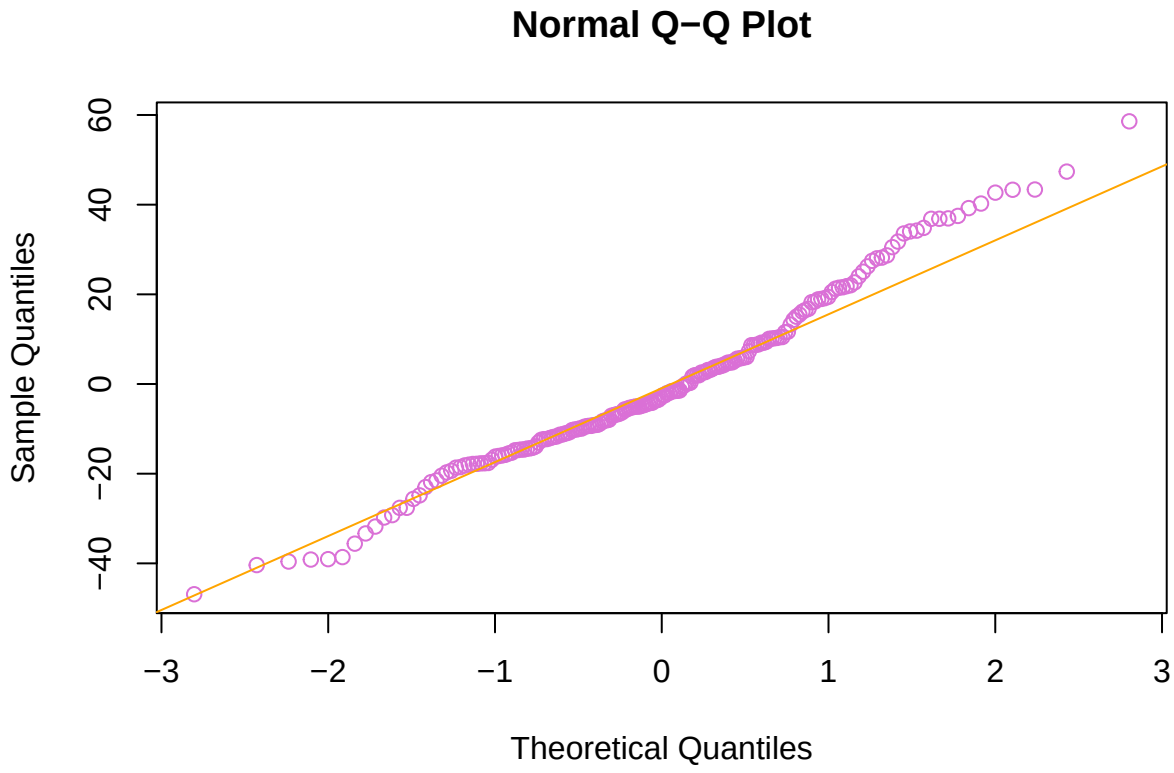
```
## Analysis of Variance Table
##
## Response: Weight
##            Df Sum Sq Mean Sq F value    Pr(>F)
## Age         1 131450  131450 357.092 < 2.2e-16
## Sex         1   8046    8046  21.858 5.488e-06
## Age:Sex     1   4368    4368  11.866 0.0007004
## Residuals 194  71414     368
```

Residuals.

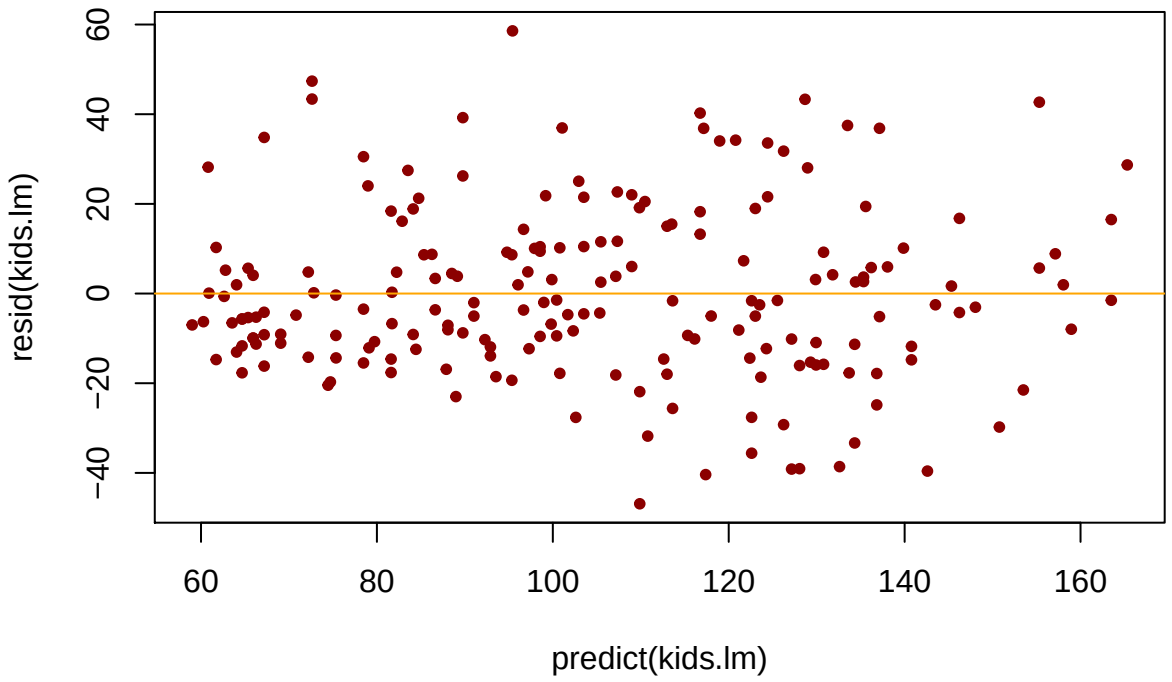
```
hist(resid(kids.lm),
     col="wheat")
```




```
qqnorm(resid(kids.lm),
       col="orchid")
qqline(resid(kids.lm),
       col="orange")
```



```r
plot(predict(kids.lm), resid(kids.lm),
     pch=20, col="darkred")
abline(h=0, col="orange")
```



CI for a difference of slopes.

```r
beta3 <- -0.28122
alpha <- 0.05
n <- 198
k <- 3
t.star <- qt(c(alpha/2, 1 - alpha/2), df=n - k - 1)
se <- 0.08164
ci <- beta3 + t.star * se
ci
```

```
## [1] -0.4422359 -0.1202041
```