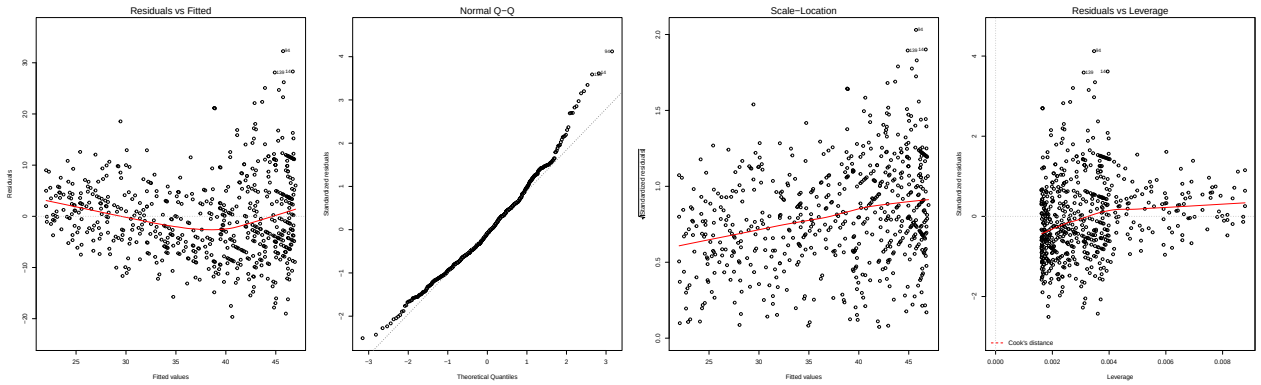
STAT 217: Transformations (in class 4/1)

1. A high respiratory rate is a potential diagnostic indicator of respiratory infection in children. To judge whether a respiratory rate is truly high, a physician must have a clear picture of the distribution of normal respiratory rates. To this end, Italian researchers measured the respiratory rates of 618 children between the ages of 15 days and 3 years.

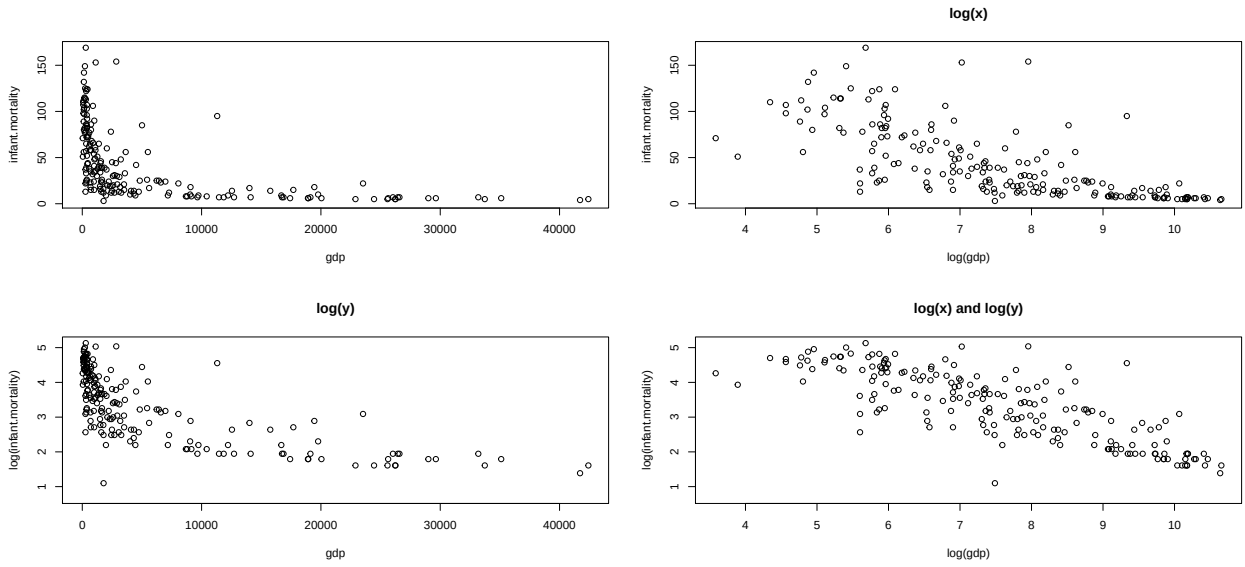
```
lm.simple <- lm(Rate~Age, data=kids)
summary(lm.simple)$coef
```

```
##             Estimate Std. Error t value  Pr(>|t|)
## (Intercept)  47.0522    0.50422   93.32 0.000e+00
## Age          -0.6957    0.02938  -23.68 1.169e-88
```

```
par(mfrow=c(1,4))
plot(lm.simple)
```



2. The UN data frame contains the infant mortality rate (infant deaths per 1000 live births) and the GDP (US dollars) of 207 countries around the world. The data are for 1998.

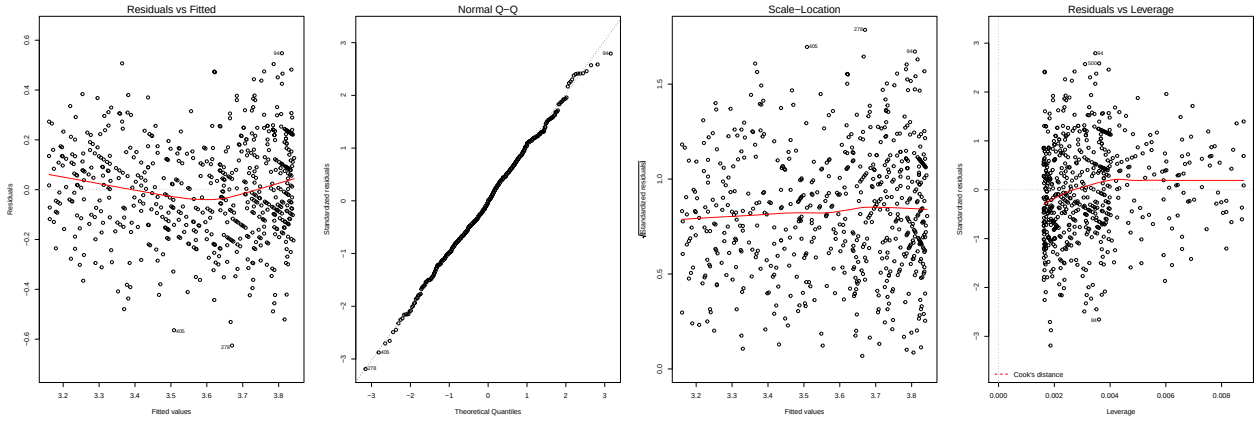


```
log.rate <- log(kids$Rate)
lm.kids <- lm(log.rate~Age, data=kids)
summary(lm.kids)$coef
```

```
##             Estimate Std. Error t value  Pr(>|t|)
## (Intercept)  3.84512  0.0126277  304.50  0.00e+00
## Age         -0.01901  0.0007357  -25.84 2.74e-100
```

```
confint(lm.kids)
```

```
##                2.5 %   97.5 %
## (Intercept)  3.82032  3.86992
## Age         -0.02045 -0.01756
```



- What is the estimated simple linear regression model before the transformation? Why is it inappropropriate to use this model for inference?

- What is the estimated simple linear regression model after the transformation?

  - Find the estimated log(rate) for children aged 28 months.

  - Find the estimated rate for children aged 28 months.