```
> AIC(model1)
[1] 5290.354
> AIC(model2)
[1] 5286.855
```

We defined `model2` as having just two predictors - as `model2` has the lower AIC value (so more information in our data is explained by `model2` than by `model1`), we would be justified in selecting that as our 'best' model.  AIC penalises models with increasing number of parameters (but not as much as BIC) so gives us a good trade-off of fitting our data and model complexity.

In regression our residuals need to be normally distributed - the easiest way to check this is to plot them:

```
> hist(residuals(model2))
```

Histogram of residuals(model2)
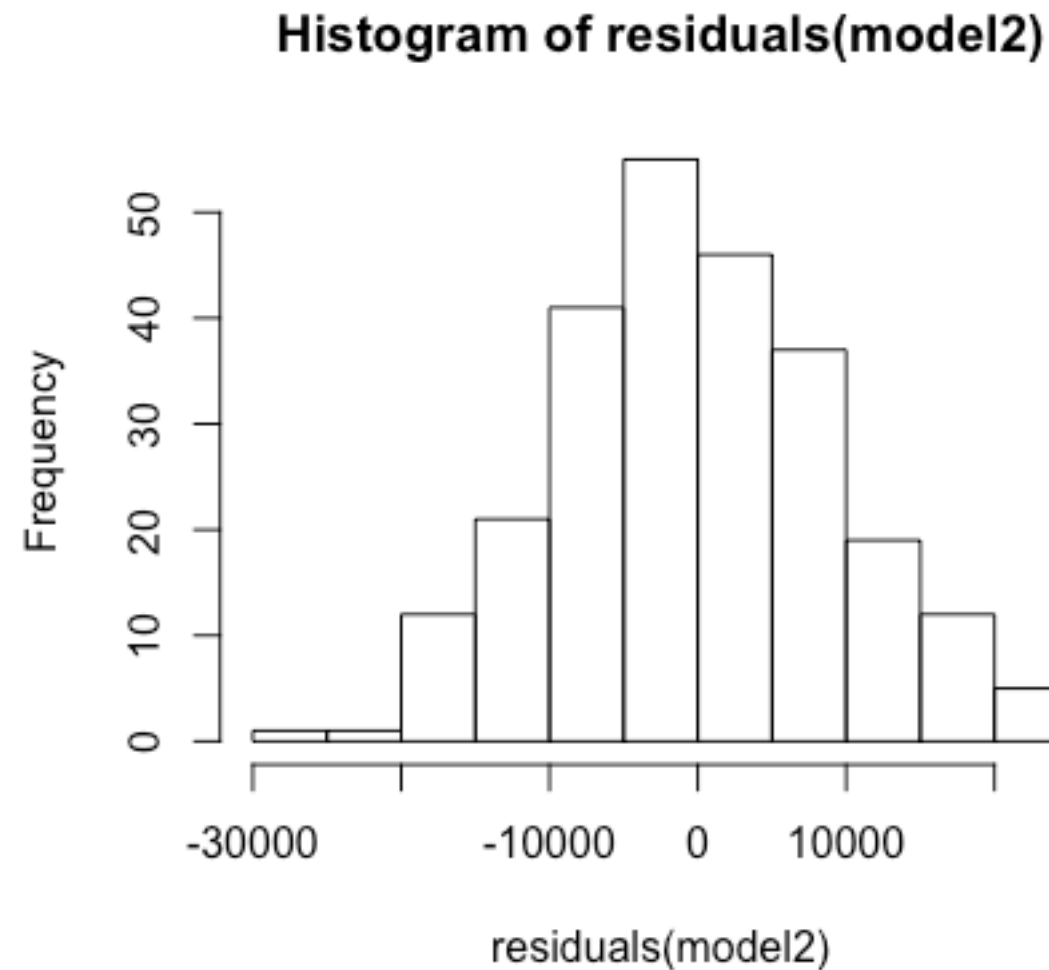
Now let's look at a number of diagnostic plots...