DAY 2 (Chap 5 ANOVA) EXERCISES SOLUTIONS

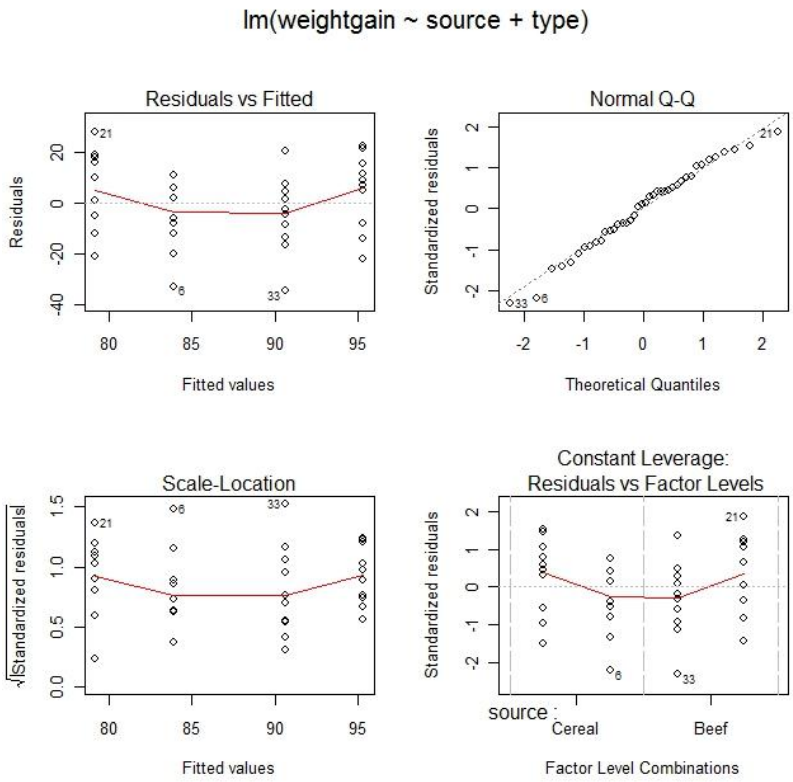
Here are my solutions to the after-class exercises from Day 2. All of the data sets referenced in these exercises may be accessed through the HSAUR2 package which is available on CRAN.

1) Examine the residuals (observed value less fitted value) from fitting a main effects only model to the **weightgain** data. What conclusions do you draw?

   There are a couple of ways to do this. You can do it by first running a linear model that looks like:

   rats_lm_1 <- lm(weightgain ~ source + type, data=weightgain)
   summary(rats_lm_1)

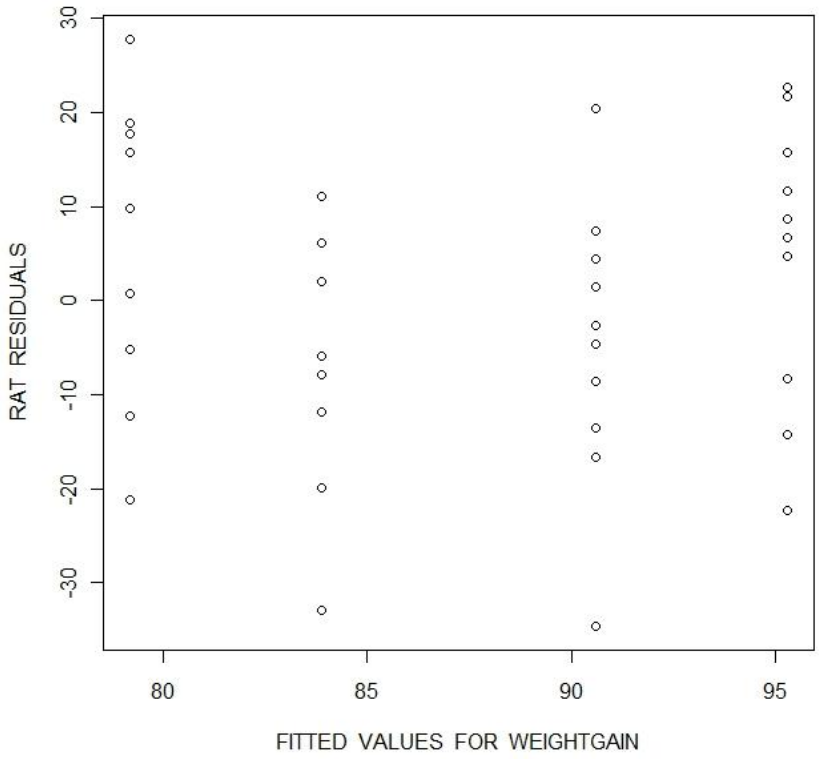
   Note: You do not need the summary statement in the above script for this exercise. Then you can either examine the residuals through the menus (Models > Graphs > Basic diagnostic plots) in which case you will see the following:



   Or, instead of using the menus, you can create a plot of the residuals versus the fitted values with your own script which might look something like:
   rat_res <- residuals(rats_lm_1)

plot(predict(rats_lm_1), rat_res, xlab="FITTED VALUES FOR WEIGHTGAIN", ylab="RAT RESIDUALS")

In which case you should see a plot which looks like this:



The two plots are the same. It appears as if the residuals are normally distributed (see QQ plot top right), are homoscedastic, and mostly have a mean near 0. The residuals do seem to dip a bit for the middle fitted values (e.g. 85 and 90), but do not appear to be grossly distorted (e.g. funnel-shaped) on either end of the range of fitted values.

2) Show how the analysis of variance table for the **weightgain** data can be constructed from the results of applying an appropriate multiple linear regression to the data.

   If you want to include the main effects and the interaction term, you can use the following script to run a linear model (same script for a regression model):

```
m1 <- lm(weightgain ~ type*source, data=weightgain)
summary(m1)
```

Then run an Anova table with this script:

```
Anova(m1, type="II")
```

You will then see the following results:

Anova Table (Type II tests)

Response: weightgain

	Sum Sq	Df	F value	Pr(>F)	
type	1299.6	1	5.8123	0.02114	*
source	220.9	1	0.9879	0.32688	
type:source	883.6	1	3.9518	0.05447	.
Residuals	8049.4	36			

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1