

## ANOVAs and LM

Remember to set your working directory as always

Write the code yourself this time

Ok now read in the data

```
MammalData <- read.csv("MammalData.csv")
```

First of all take a look at the data using str, head and summary. Maybe even view it.

```
str(MammalData)
```

etc

While the numbers being spat out at us might be useful it could be useful to plot the data.

Considering one of our assumptions is normality one of the most useful plots we can do is a histogram.

histograms are plotted with a simple command

```
hist(MammalData$meanCvalue)
```

fill in the data you want plotted.

After taking a look at our plots are we happy that it is normal?

look at meanCvalue

Normal enough at least....

So we have a good idea of what our data looks like now it's time to do an ANOVA.

What are we interested to know? Well how about if the order affects the Cvalue

ok so how do we put this into our ANOVA

```
aov(meanCvalue ~ Order, data = MammalData)
```

when we put things against each other in r we use ~ this is the same for models and plots.

That spat out a lot of numbers at us and we will have to rerun it again and again whenever we want them.

How about we save our model as an object?

```
ano1 <- aov(meanCvalue ~ Order, data = MammalData)
```

now we can see it whenever we call the object and do lots of things with it.

```
summary(ano1)
```

that gave us pretty much what we got before and is most of the specifics we need for a test

```
str(ano1)
```

this gives us a whole tonne of different things but is particularly useful if we want to know the coordinates of things inside the anova.

Now we've done an ANOVA lets look at a LM.

lets look at the very simple dataset called trees

there is an premade dataset in R called trees

```
trees <- trees
```

take a look at the hist of girth and height

those two look kind of normal lets see if they are related

```
lm(Girth ~ Height, data = trees)
```

again it spits it out, remeber to make it an object

```
m1 <- lm(Girth ~ Height, data = trees)
```

```
summary(m1)
```

But how well did our model fit the assumptions that we talked about?

```
plot(m1)
```

if we use the plot functions it makes 4 seperate plots looking at each of the assumptions.

Take a look through the plots and see if you can identify which one links to each assumption

this model really doesn't seem that bad

lets test to see if volume is related to girth

```
m2 <- lm(Volume ~ Girth, trees)
```

```
summary(m2)
```

```
plot(m3)
```

Ok girth is related to Volume, how about height. Or even better see if they work together on volume

```
m3 <- lm(Volume ~ Girth * Height, trees)
```

When we want to see if things interact in a model we use \* when we just want to put them both in we use +

```
summary(m3)
```

```
plot(m3)
```

In this case everything is significant which means both girth and height have an effect on volume and they interact meaning changes in height also affect changes in girth etc.

Ok Now I want you to open the ocean carbon dataset and investigate it. Find out the relationship between carbon sequestration rate and ocean temperature.  
tell me if year and month are having an effect.

hint year is a variable that should be a factor turn it into one by using `as.factor(Year)`  
or if you want to be very clever you can try and use it as a time