Looking down the list of parameter estimates, we see that the most similar are the effects of root pruning to 10 and 5 cm (145.5 vs. 145.33). We shall begin by simplifying these to a single root-pruning treatment called root. The trick is to use the gets arrow to change the names of the appropriate factor levels. Start by copying the original factor name:

clip2<-clipping

Now inspect the level numbers of the various factor level names:

levels(clip2)

[1] "control"  "n25"  "n50"  "r10"  "r5"

The plan is to lump together r10 and r5 under the same name, ‘root’. These are the fourth and fifth levels of clip2, so we write:

levels(clip2)[4:5]<-"root"

If we type

levels(clip2)

[1] "control"  "n25"  "n50"  "root"

we see that r10 and r5 have indeed been replaced by root.

The next step is to fit a new model with clip2 in place of clipping, and to test whether the new simpler model is significantly worse as a description of the data using anova:

model4<-aov(biomass ~ clip2)
anova(model3,model4)

Analysis of Variance Table
Model 1: biomass~clipping
Model 2: biomass~clip2

<table>
<thead>
<tr>
<th>Res.Df</th>
<th>RSS</th>
<th>Df</th>
<th>Sum of Sq</th>
<th>F</th>
<th>Pr(&gt;F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>124020</td>
<td>-1</td>
<td>-0.0833333</td>
<td>0.9968</td>
<td></td>
</tr>
</tbody>
</table>

As we expected, this model simplification was completely justified.

The next step is to investigate the effects using summary.lm:

summary.lm(model4)

Coefficients:

| Estimate | Std. Error | t value | Pr(>|t|) |
|----------|------------|---------|----------|
| (Intercept) | 465.17     | 28.20   | 16.498   | 2.66e-15 *** |
| clip2n25   | 88.17      | 39.87   | 2.211    | 0.036029 *   |
| clip2n50   | 104.17     | 39.87   | 2.612    | 0.014744 *    |
| clip2root  | 145.42     | 34.53   | 4.211    | 0.000269 ***  |

It looks as if the two shoot-clipping treatments (n25 and n50) are not significantly different from one another (they differ by just 16.0 with a standard error of 39.87). We can lump these together into a single shoot-pruning treatment as follows:

clip3<-clip2
levels(clip3)[2:3]<-"shoot"
levels(clip3)

[1] "control"  "shoot"  "root"
Then fit a new model with clip3 in place of clip2:

```r
test <- aov(biomass ~ clip3)
anova(test)
```

Analysis of Variance Table
Model 1: biomass ~ clip2
Model 2: biomass ~ clip3

<table>
<thead>
<tr>
<th>Res.Df</th>
<th>RSS</th>
<th>Df</th>
<th>Sum of Sq</th>
<th>F Pr(&gt;F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>124020</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>124788</td>
<td>-1</td>
<td>-768</td>
<td>0.161</td>
</tr>
</tbody>
</table>

Again, this simplification was fully justified. Do the root and shoot competition treatments differ?

```r
clip4 <- clip3
levels(clip4)[2:3] <- "pruned"
levels(clip4)
```

[1] "control" "pruned"

Now fit a new model with clip4 in place of clip3:

```r
test <- aov(biomass ~ clip4)
anova(test)
```

Analysis of Variance Table
Model 1: biomass ~ clip3
Model 2: biomass ~ clip4

<table>
<thead>
<tr>
<th>Res.Df</th>
<th>RSS</th>
<th>Df</th>
<th>Sum of Sq</th>
<th>F Pr(&gt;F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>124788</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>139342</td>
<td>-1</td>
<td>-14553</td>
<td>3.1489</td>
</tr>
</tbody>
</table>

This simplification was close to significant, but we are ruthless ($p > 0.05$), so we accept the simplification. Now we have the minimal adequate model:

```r
summary.lm(test)
```

Coefficients:
```
            Estimate Std. Error t value Pr(>|t|)  
(Intercept) 465.2       28.8     16.152 1.11e-15 ***  
clip4pruned 120.8       32.2      3.751   0.000815 ***  
```

It has just two parameters: the mean for the controls (465.2) and the difference between the control mean and the 4 treatment means (465.2 + 120.8 = 586.0):

```r
tapply(biomass, clip4, mean)
```

```
control  pruned  
465.1667  585.9583  
```

We know that these two means are significantly different from the $p$ value of 0.000 815, but just to show how it is done, we can make a final model7 that has no explanatory variable at all (it fits only the overall mean). This is achieved by writing $y ~ 1$ in the model formula:

```r
test7 <- aov(biomass ~ 1)
anova(test, test7)
```