NRES_798_9_201501

Linear model

applications

Fixed vs. Random factors

- Two different types of categorical explanatory variables

- Fixed
  - Treatment levels are the only ones of interest
  - Inference restricted to the treatment levels
  - Treatment contains constant information that is universal across populations (e.g. sex)
  - Set or selected by researcher

- Random
  - Treatment levels are a random sample of all possible levels
  - Inferences intended to extend beyond treatment levels tested
  - Treatments are not constrained and constant across populations (e.g. genotype)
  - Set by drawing samples from a broader underlying distribution

Fixed vs. Random effects

- Drug administered or not
- Block within a field
- Insecticide sprayed
- Brood
- Nutrients added
- One country vs. another
- History of development
- Split plot within a plot
- Family
- Untreated individuals
- Parent
- Light vs. shade
- One age vs. another

Fixed vs. Random effects (blue fixed)

- Drug administered or not
- Block within a field
- Insecticide sprayed
- Brood
- Nutrients added
- One country vs. another
- History of development
- Split plot within a plot
- Family
- Untreated individuals
- Parent
- Light vs. shade
- One age vs. another

Fixed vs. Random factor

- Mean square values and F-ratio structure differ between fixed and random factors
  - Only applicable to two-way ANOVA

- Fixed factor
  - F-ratio: main effects and interactions are compared against the residual mean square

$$F - ratio_A = \frac{explained\ variance}{unexplained\ variance} = \frac{MS_A}{MS_{within\ groups}}$$

$$F - ratio_{AB} = \frac{explained\ variance}{unexplained\ variance} = \frac{MS_{AB}}{MS_{within\ groups}}$$

  - Focus on increasing replication within treatments
    - i.e. inference focuses on treatments therefore replicates increase power

Fixed vs. Random factor

- Random factor
  - F-ratio: main effects are tested against the interaction mean square, interaction effect tested against within group mean square

$$F - ratio_A = \frac{explained\ variance}{unexplained\ variance} = \frac{MS_A}{MS_{AB}}$$

$$F - ratio_{AB} = \frac{explained\ variance}{unexplained\ variance} = \frac{MS_{AB}}{MS_{within\ groups}}$$

  - Focus on increasing the number of treatment levels established
    - i.e. inference focuses on sampling distribution therefor more treatment levels increase power
    - When to give up on ANOVA and use regression?
  - Random block and repeated measures analysed as random factors

Two-way ANOVA

ANIMAL BEHAVIOUR, 2000, 60, 511–521
doi:10.1006/anbe.2000.1513, available online at http://www.idealibrary.com on IDEAL®

Lack of preference for low-predation-risk habitats in larval damselflies explained by costs of intraspecific interactions

CHÉ M. ELKIN & ROBERT L. BAKER
Department of Zoology, University of Toronto at Mississauga

(Received 16 August 1998; initial acceptance 28 September 1999; final acceptance 26 April 2000; MS. number: A8566)

Many studies indicate prey organisms select microhabitats with high structural complexity as a way of reducing risk of predation. We used laboratory experiments to show that damselfly larvae, *Ischnura verticalis*, suffer higher predation rates from pumpkinseed sunfish in low-density vegetation. However, larvae do not preferentially occupy microhabitats with high vegetation density in either the presence or absence of sunfish; when given a choice, the number of larvae per stem of vegetation was equal across all densities of vegetation. That larvae do not congregate in dense vegetation may reflect costs of aggressive interactions. Results from laboratory experiments indicated larval interactions increase conspicuous behaviours (most notably swimming) and consequently increase fish predation. A subsequent experiment indicated that frequency of larval interactions increases with increased vegetation density when number of larvae/stem is constant. Thus, larval microhabitat selection may reflect a trade-off between reduced risk of predation in areas of high vegetation density, caused by reduced fish foraging ability, and increased aggressive larval interactions, due to decreased proximity of larvae.

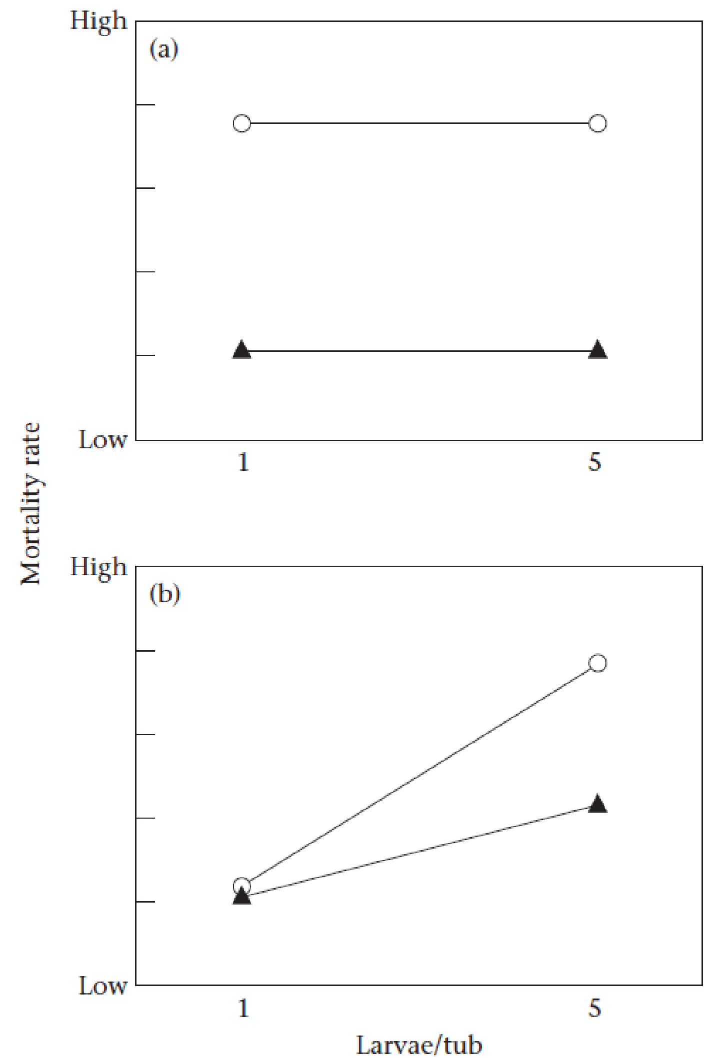


Figure 3. Scenarios of possible mortality rates. (a) Expected mortality due to an increase in the ability of fish to detect larvae, independent of larval interactions. (b) Expected mortality as positively related to frequency of larval interactions. (○: low vegetation; ▲: high vegetation) See text for details.

Two-way ANOVA

- Vegetation density (p = 0.0002)
- Conspecifics (p = 0.0026)
- Interaction (p = 0.1589)

- Log10(X +1) transformed

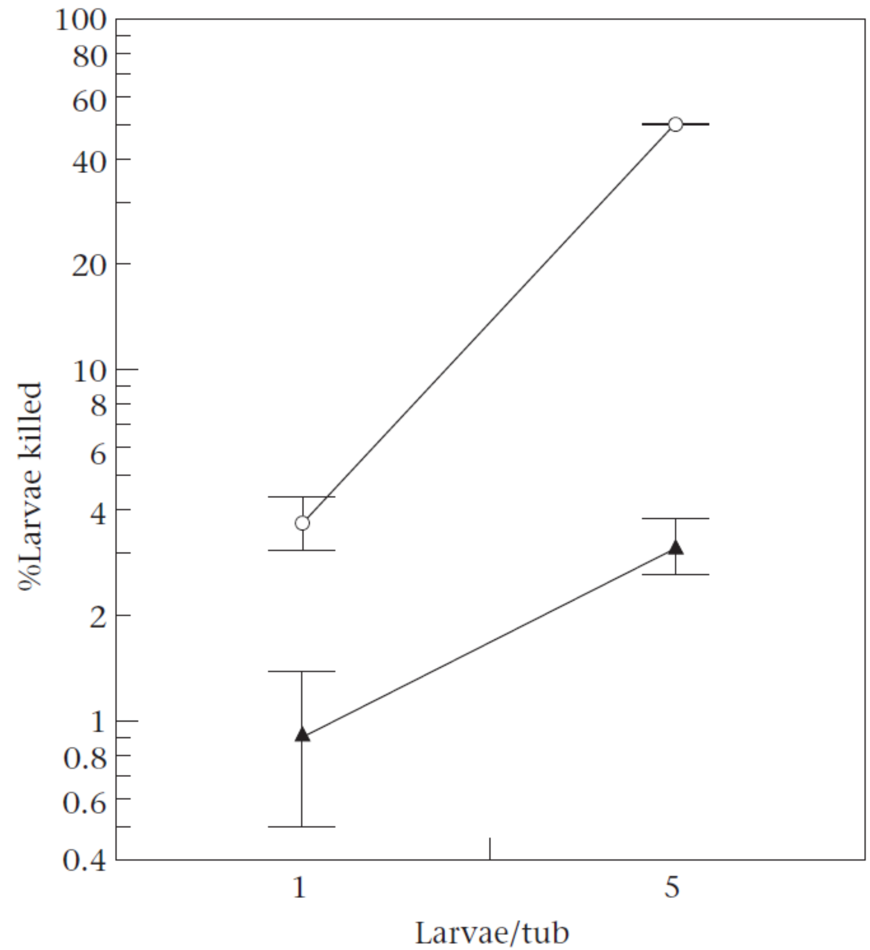
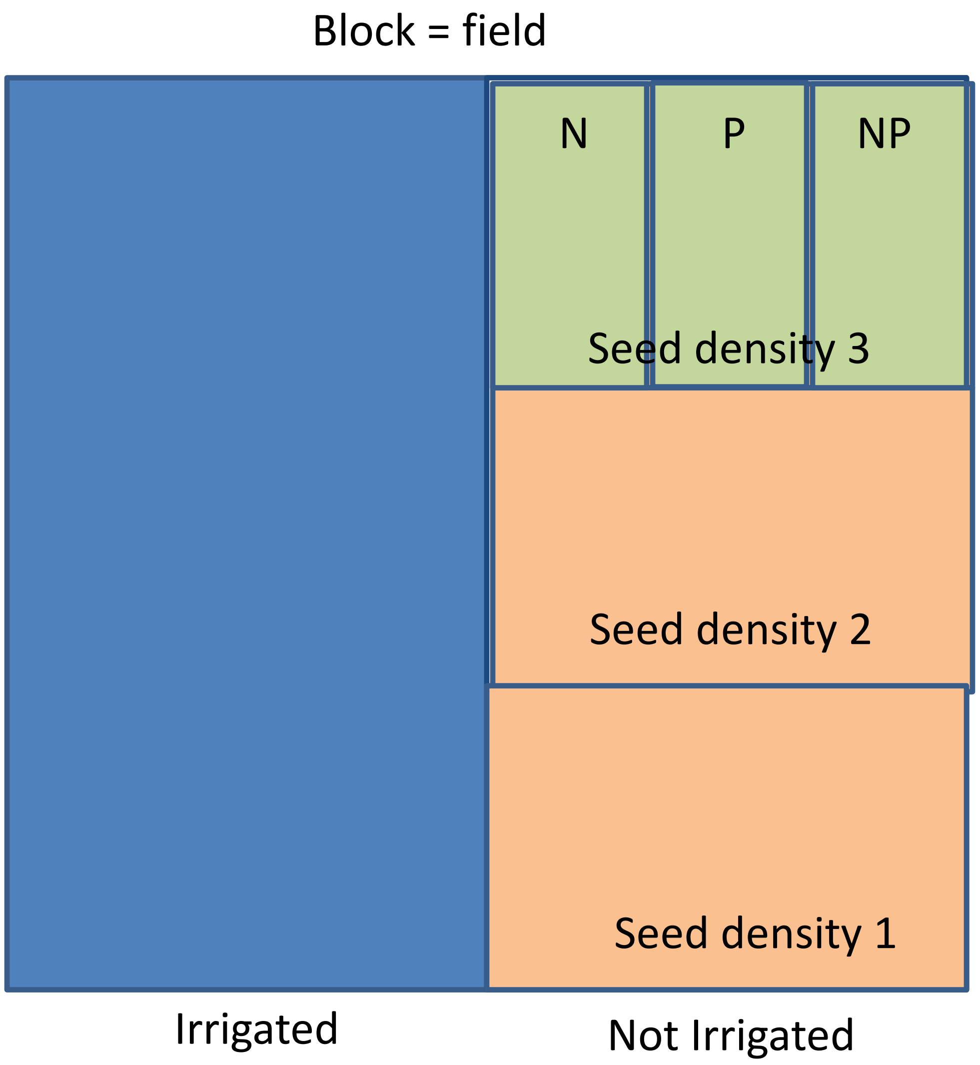


Figure 4. Mean±SE percentage of larvae killed under various conditions of larvae (one versus five per tub) and vegetation density. (○: low vegetation; ▲: high vegetation) Values presented are back transformed.

Split plot ANOVA

- Experiment examining the influence of irrigation, seed density and fertilization on crop yield

- Blocks: 4 fields
- Irrigation: 2 levels
- Seeding density: 3 levels
- Fertilization: 3 levels



Split plot ANOVA

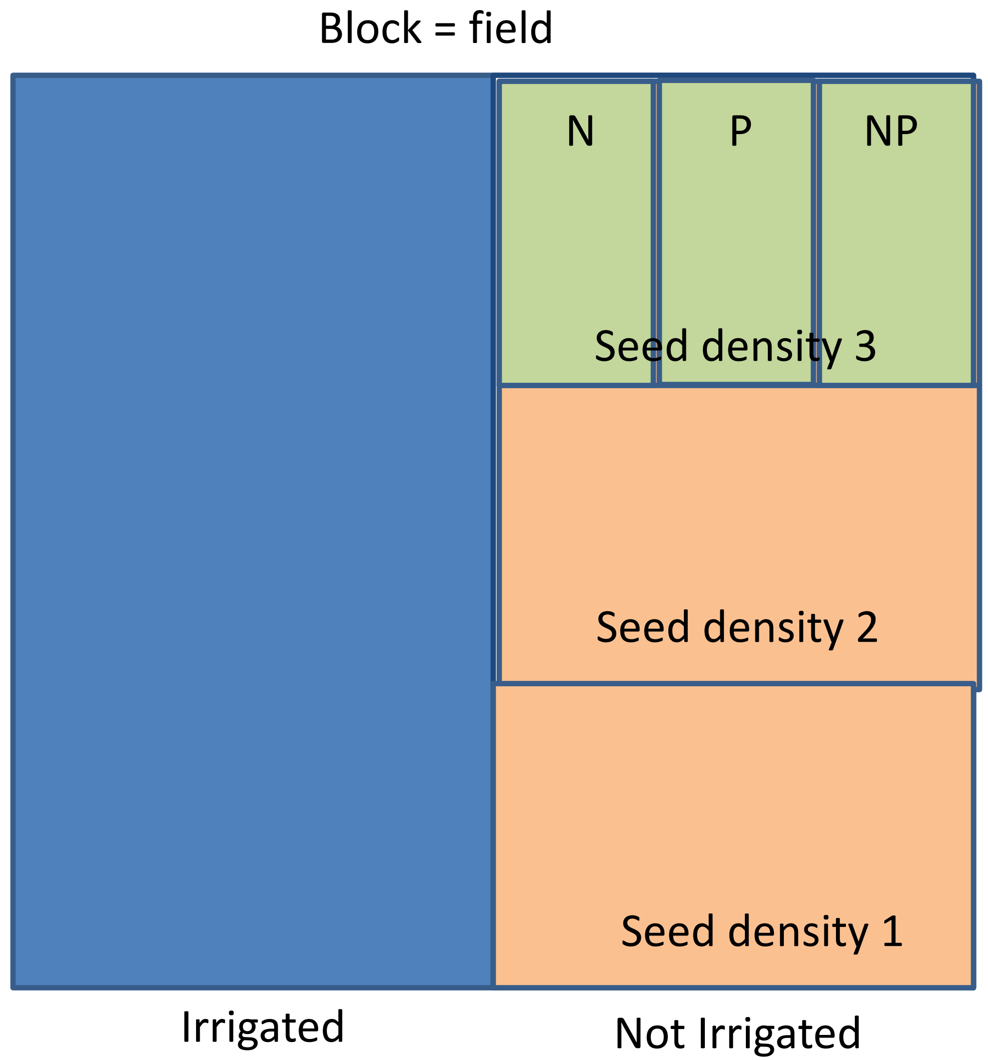
```
> str(yields)
'data.frame':    72 obs. of  5 variables:
 $ yield     : int  90 95 107 92 89 92 81 92 93 80 ...
 $ block     : Factor w/ 4 levels "A","B","C","D": 1 1 1 1 1 1 1 1 1 1 ...
 $ irrigation: Factor w/ 2 levels "control","irrigated": 1 1 1 1 1 1 1 1 1 2 ...
 $ density   : Factor w/ 3 levels "high","low","medium": 2 2 2 3 3 3 1 1 1 2 ...
 $ fertilizer: Factor w/ 3 levels "N","NP","P": 1 3 2 1 3 2 1 3 2 1 ...
```

4 *2 * 3 * 2 = 72

```
> yields
   yield block irrigation density fertilizer
1     90     A    control     low          N
2     95     A    control     low          P
3    107     A    control     low         NP
4     92     A    control  medium          N
5     89     A    control  medium          P
6     92     A    control  medium         NP
7     81     A    control    high          N
8     92     A    control    high          P
9     93     A    control    high         NP
10    80     A  irrigated     low          N
11    87     A  irrigated     low          P
```

Split plot ANOVA

```
model_bad <- aov(yield~irrigation*density*fertilizer)
summary(model_bad)
```



Split plot ANOVA

```
model_bad <- aov(yield~irrigation*density*fertilizer)
summary(model_bad)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
irrigation	1	8278	8278	59.575	2.81e-10	***
density	2	1758	879	6.328	0.00340	**
fertilizer	2	1977	989	7.116	0.00181	**
irrigation:density	2	2747	1374	9.885	0.00022	***
irrigation:fertilizer	2	953	477	3.431	0.03956	*
density:fertilizer	4	305	76	0.549	0.70082	
irrigation:density:fertilizer	4	235	59	0.422	0.79183	
Residuals	54	7503	139			

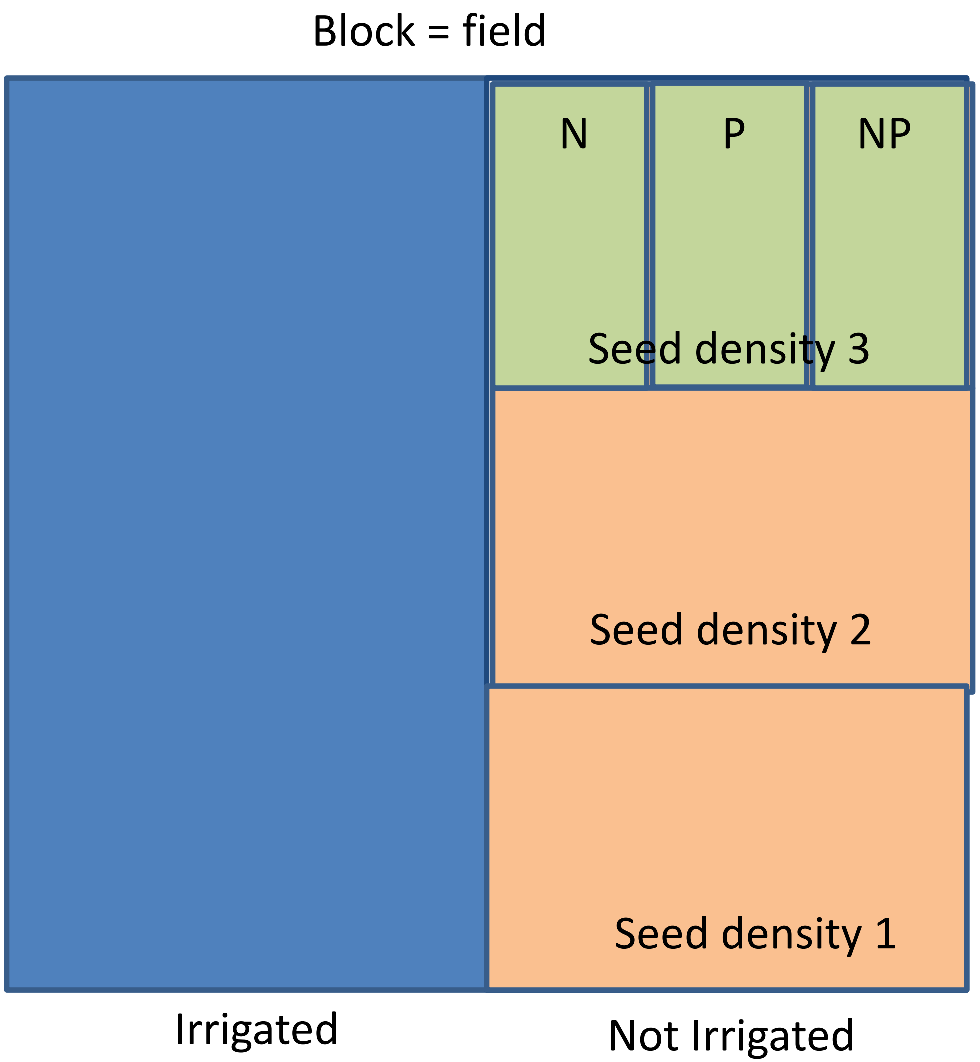
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Split plot ANOVA

model <- aov(yield ~ irrigation*density*fertilizer + Error(block/irrigation/density))
summary(model)

Four ANOVA tables one for each plot size

```
Error: block
                  Df           Sum Sq Mean Sq F value Pr(>F)
Residuals         3            194.4   64.81

Error: block:irrigation
                  Df           Sum Sq Mean Sq F value Pr(>F)
irrigation        1            8278    8278   17.59 0.0247 *
Residuals         3            1412     471
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Error: block:irrigation:density
                  Df           Sum Sq Mean Sq F value Pr(>F)
density           2            1758    879.2   3.784   0.0532 .
irrigation:density  2          2747    1373.5   5.912  0.0163 *
Residuals         12           2788    232.3
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Error: Within
                             Df  Sum Sq Mean Sq F value   Pr(>F)
fertilizer                    2  1977.4  988.7  11.449 0.000142 ***
irrigation:fertilizer         2  953.4   476.7   5.520 0.008108 **
density:fertilizer            4  304.9   76.2    0.883 0.484053
irrigation:density:fertilizer 4  234.7   58.7    0.680 0.610667
Residuals                    36  3108.8   86.4
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```



Split plot ANOVA

```
model <- aov(yield ~ irrigation*density*fertilizer + Error(block/irrigation/density))
summary(model)
```

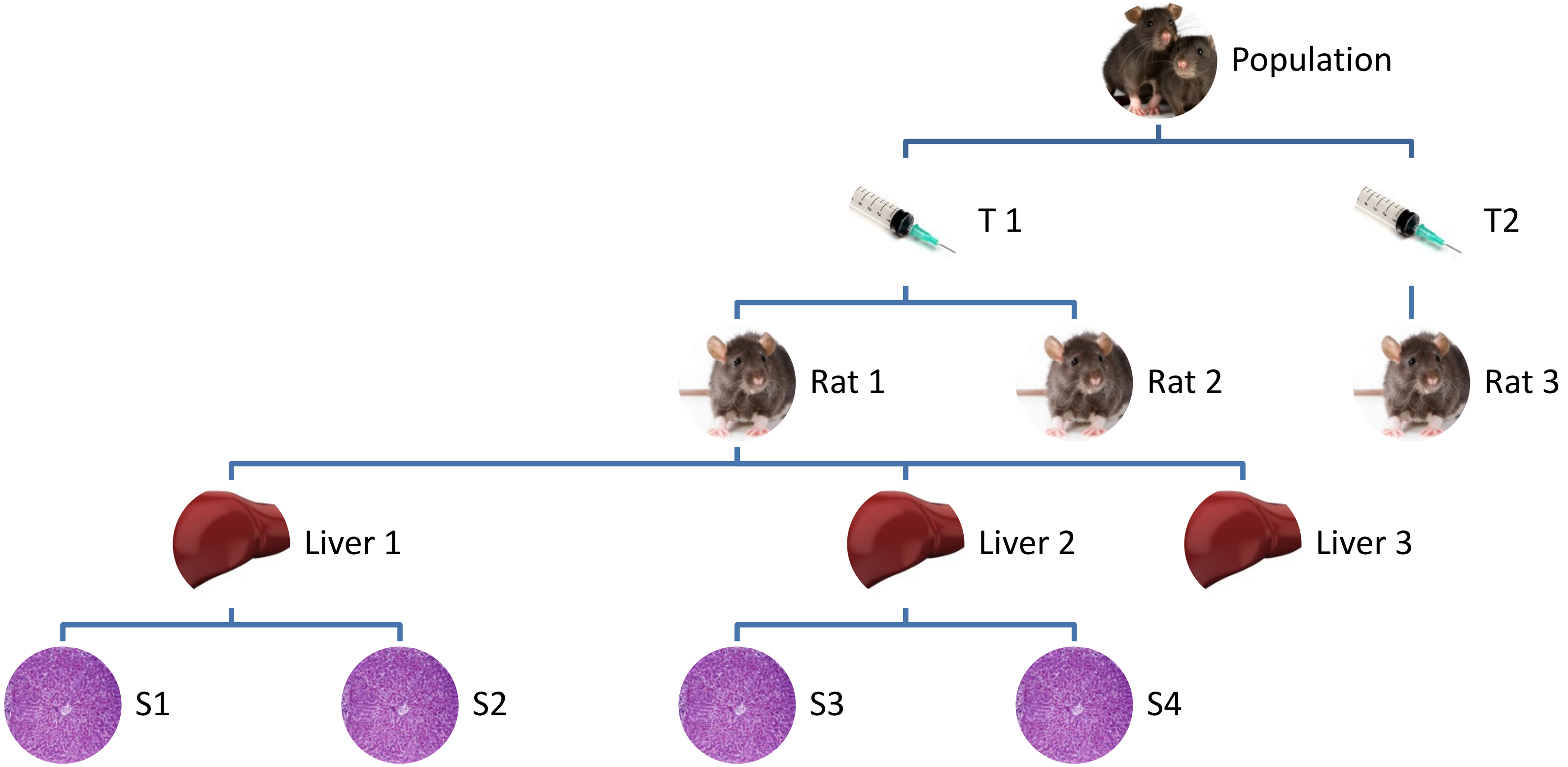
```
Error: block
          Df Sum Sq Mean Sq F value Pr(>F)
Residuals  3  194.4   64.81

Error: block:irrigation
                Df Sum Sq Mean Sq F value Pr(>F)
irrigation       1   8278    8278   17.59 0.0247 *
Residuals        3   1412     471
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Error: block:irrigation:density
                   Df Sum Sq Mean Sq F value Pr(>F)
density             2   1758   879.2   3.784 0.0532 .
irrigation:density  2   2747  1373.5   5.912 0.0163 *
Residuals          12   2788   232.3
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Error: Within
                              Df S          um Sq Mean Sq F value   Pr(>F)
fertilizer                    2             1977.4   988.7  11.449 0.000142 ***
irrigation:fertilizer         2              953.4   476.7   5.520 0.008108 **
density:fertilizer            4              304.9    76.2   0.883 0.484053
irrigation:density:fertilizer 4              234.7    58.7   0.680 0.610667
Residuals                     36            3108.8    86.4
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
                              Df   Sum Sq Mean Sq F value   Pr(>F)
irrigation                     1   8278    8278    59.575 2.81e-10 ***
density                        2   1758     879     6.328  0.00340 **
fertilizer                     2   1977     989     7.116  0.00181 **
irrigation:density             2   2747    1374     9.885  0.00022 ***
irrigation:fertilizer          2   953      477     3.431  0.03956 *
density:fertilizer             4   305       76     0.549  0.70082
irrigation:density:fertilizer  4   235       59     0.422  0.79183
Residuals                     54   7503     139
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Nested ANOVA

- Influence of treatment on rat liver Glycogen content

3 treatments (T1, T2, T3)
2 rats/treatment
3 liver sections
2 preparations of each liver section



Nested ANOVA

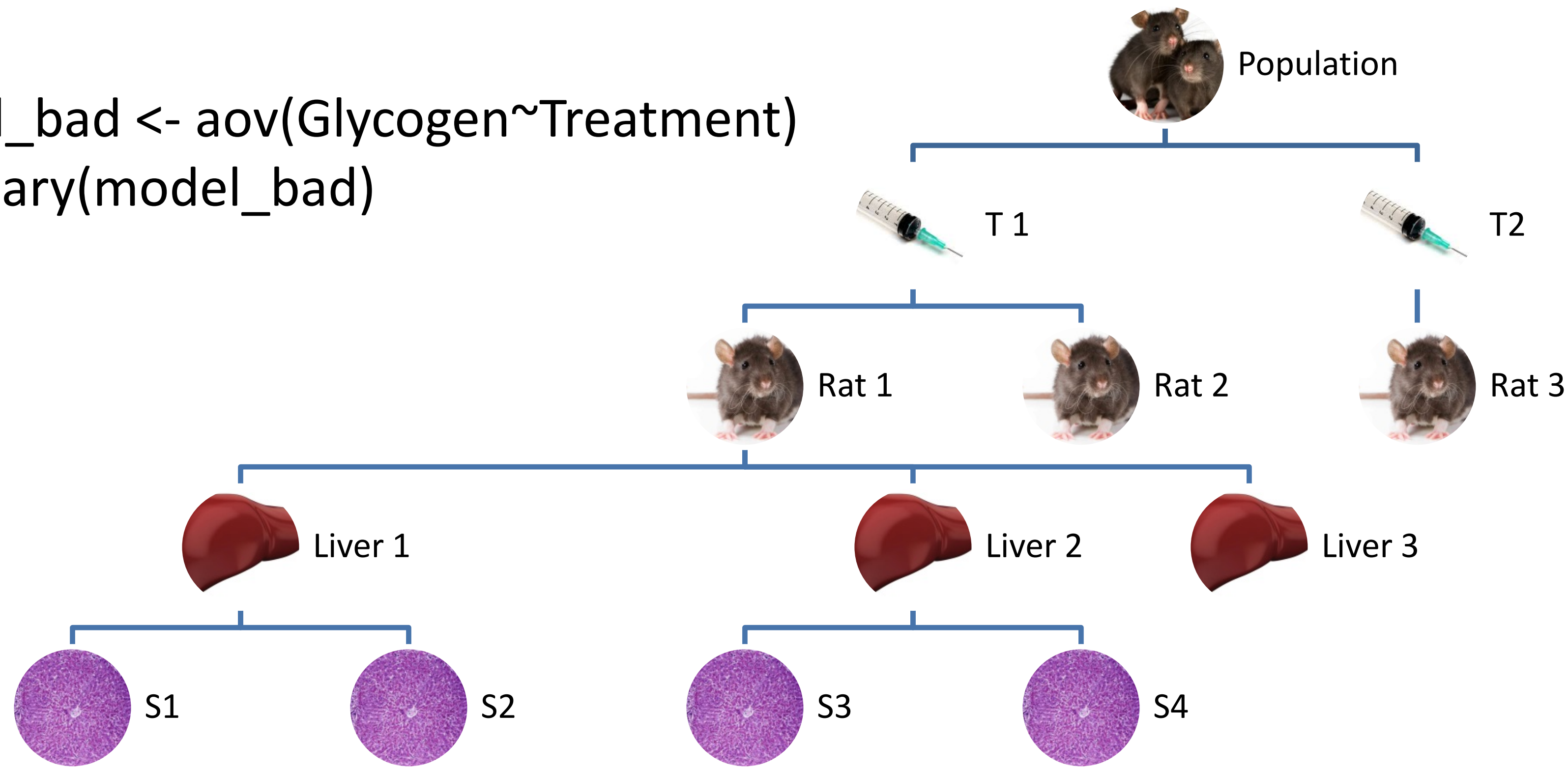
```
datpath <- "C:/Users/Che/UNBC_work/Courses/NRES-798/NRES-798-Labs/therbook/rats.txt"
rats <- read.table(datpath,header=TRUE)
attach(rats)
names(rats)
str(rats)

Treatment<-factor(Treatment)
Rat<-factor(Rat)
Liver<-factor(Liver)
```

```
> str(rats)
'data.frame':	36 obs. of  4 variables:
 $ Glycogen : int  131 130 131 125 136 142 150
148 140 143 ...
 $ Treatment: int  1 1 1 1 1 1 1 1 1 1 ...
 $ Rat      : int  1 1 1 1 1 1 2 2 2 2 ...
 $ Liver    : int  1 1 2 2 3 3 1 1 2 2 ...
```

	Glycogen	Treatment	Rat	Liver
1	131	1	1	1
2	130	1	1	1
3	131	1	1	2
4	125	1	1	2
5	136	1	1	3
6	142	1	1	3
7	150	1	2	1
8	148	1	2	1

Nested ANOVA

```
model_bad <- aov(Glycogen~Treatment)
summary(model_bad)
```



	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Treatment	1	170.7	170.67	1.837	0.184
Residuals	34	3159.6	92.93		

Pseudo replication
3*2*3*2 = 36

Nested ANOVA

```
model <- aov(Glycogen ~ Treatment + Error(Treatment/Rat/Liver))
summary(model)
```

```
Error: Treatment
          Df Sum Sq  Mean Sq
Treatment  1   170.7   170.7

Error: Treatment:Rat
          Df Sum Sq Mean Sq F value Pr(>F)
Residuals  1   148.6    148.6

Error: Treatment:Rat:Liver
          Df  Sum Sq  Mean Sq F value Pr(>F)
Residuals  1   0.03214  0.03214

Error: Within
          Df Sum Sq Mean Sq F value Pr(>F)
Residuals 32   3011   94.09
```

$$SS_{among\ groups} = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(\bar{Y}_i - \bar{Y})^2$$

$$SS_{replicates(groups)} = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}\left(\bar{Y}_{j(i)} - \bar{Y}_i\right)^2$$

$$SS_{subsamples} = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}\left(Y_{ijk} - \bar{Y}_{j(i)}\right)^2$$

The correction factor at any level is the uncorrected sum of squares from the level above

Nested ANOVA vs. Split plot ANOVA

- Nested ANOVA: uninformative factor levels
  - E.g. rat, liver bit, liver preparation
- Split plot ANOVA: informative factor levels
  - E.g. seed density, fertilizer, irrigation

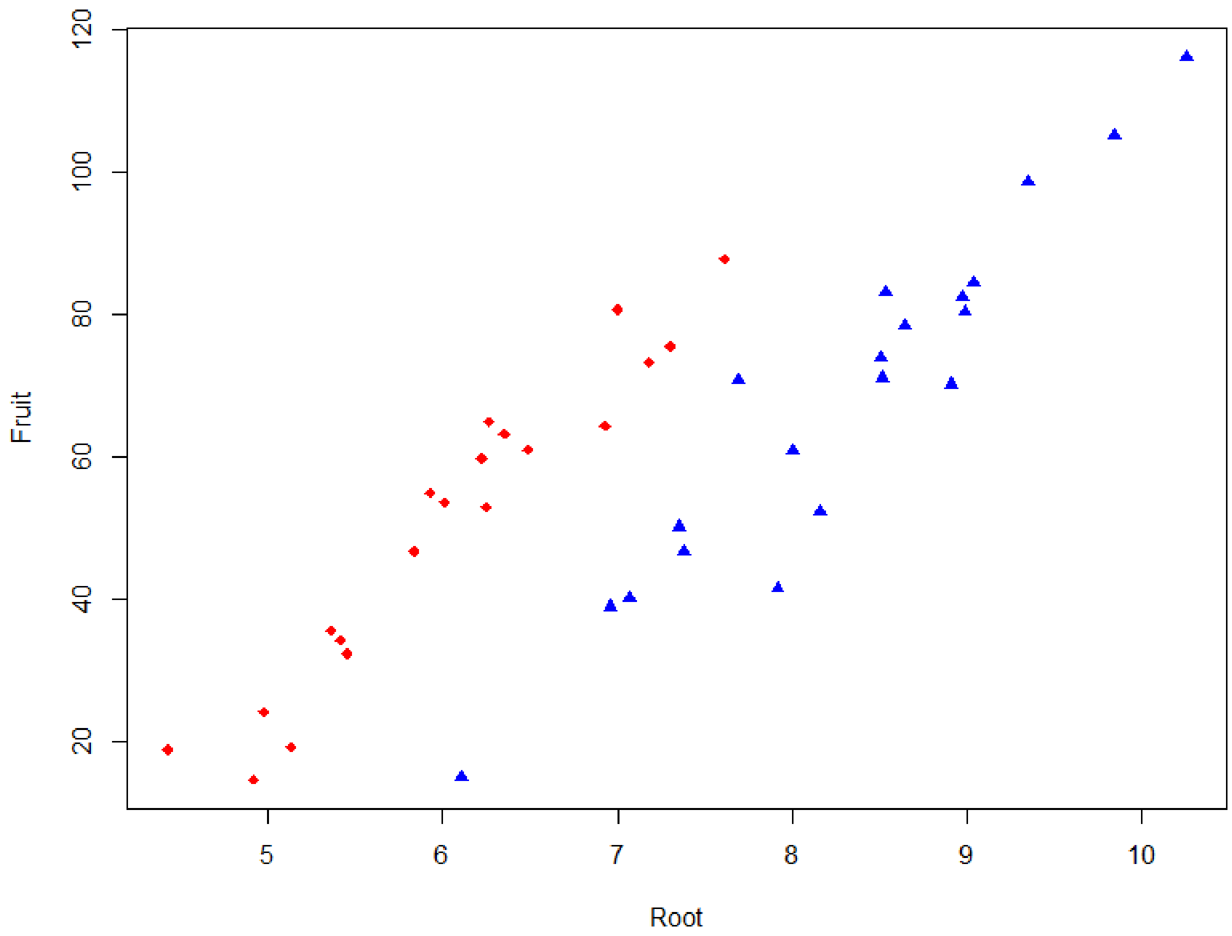
ANCOVA

- Impact of grazing on seed production of a biennial plant

- 2 treatments: grazed, ungrazed
- Diameter of rootstock measured
- Mass of fruit measured
- 40 individuals total

Ipomopsis rubra

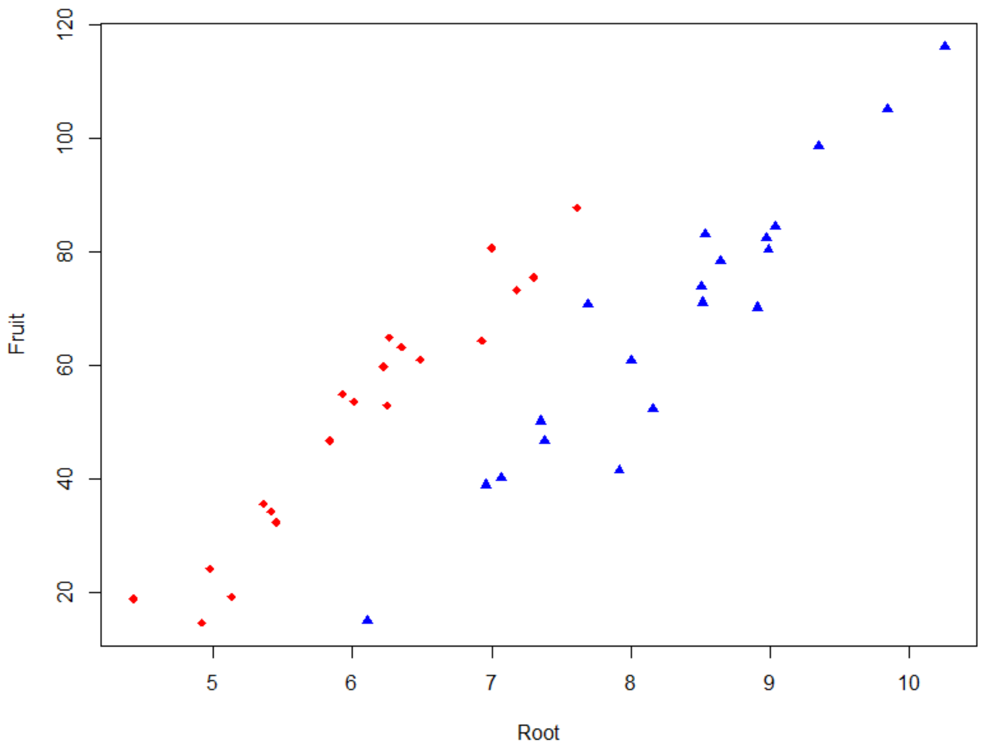
ANCOVA



- Red grazed
- Blue Ungrazed

ANCOVA

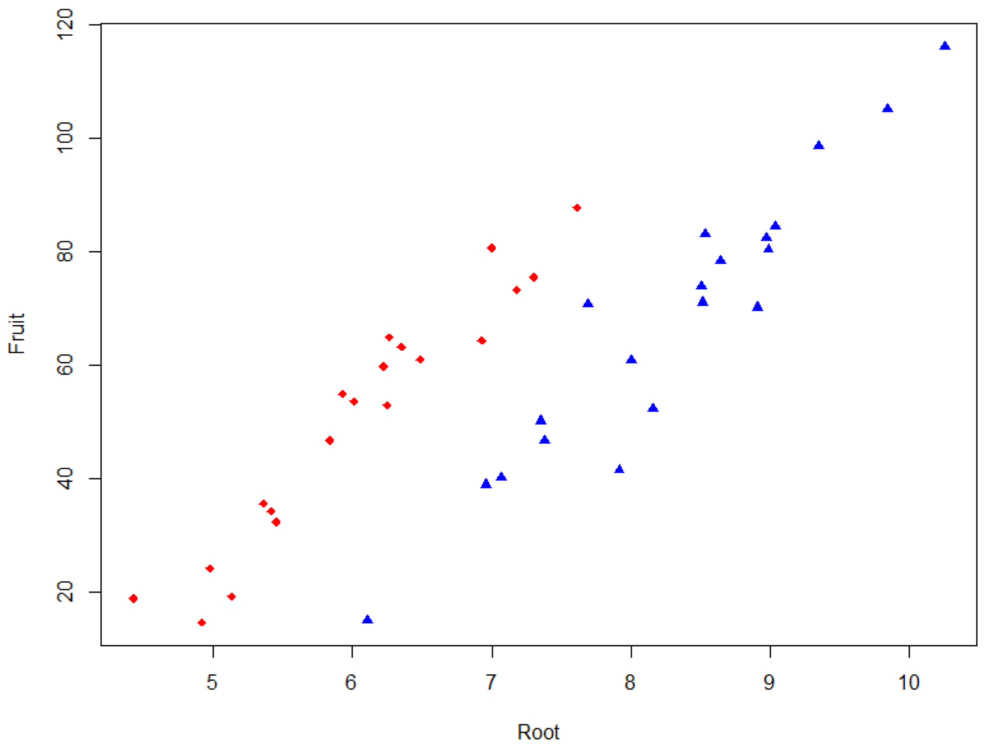
- Red grazed
- Blue Ungrazed



- Fit complex model first

- ancova <- lm(Fruit~Grazing*Root)
- summary(ancova)

ANCOVA



Call:
lm(formula = Fruit ~ Grazing * Root)

Residuals:

Min	1Q	Median	3Q	Max
-17.3177	-2.8320	0.1247	3.8511	17.1313

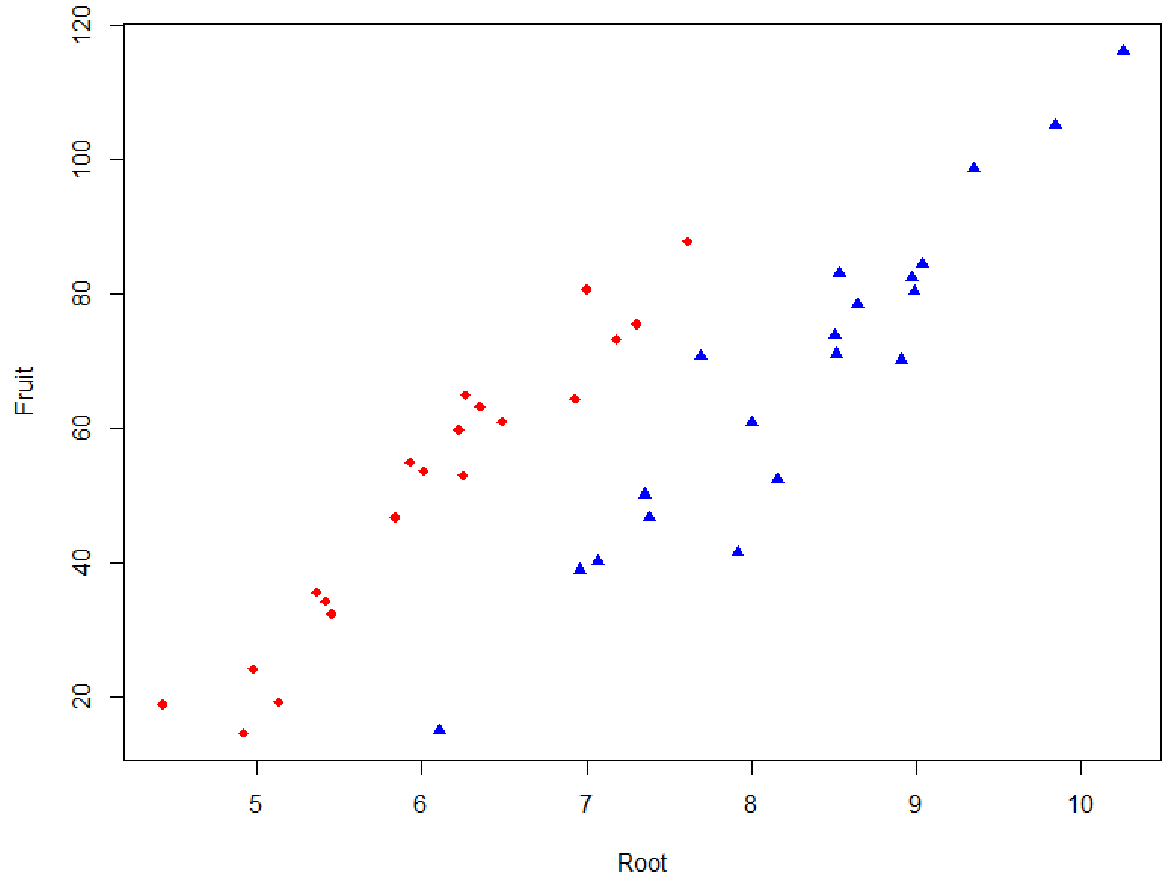
Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-125.173	12.811	-9.771	1.15e-11 ***
GrazingUngrazed	30.806	16.842	1.829	0.0757 .
Root	23.240	1.531	15.182	< 2e-16 ***
GrazingUngrazed:Root	0.756	2.354	0.321	0.7500

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 6.831 on 36 degrees of freedom
Multiple R-squared: 0.9293, Adjusted R-squared: 0.9234
F-statistic: 157.6 on 3 and 36 DF, p-value: < 2.2e-16

ANCOVA



Analysis of Variance Table

Response: Fruit

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Grazing	1	2910.4	2910.4	62.3795	2.262e-09 ***
Root	1	19148.9	19148.9	410.4201	< 2.2e-16 ***
Grazing:Root	1	4.8	4.8	0.1031	0.75
Residuals	36	1679.6	46.7		

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1