Figures

```r
library(tidyverse)
library(smmr)
library(broom)
```

Figure 1: Packages

```
species weight
     RT    920
     RT    945
     CH    330
     RT    970
     RT   1240
     RT   1290
     CH    470
     CH    335
     RT   1150
     RT    960
     RT   1130
     RT    925
     RT   1205
     RT   1040
     CH    335
     CH    335
```

Figure 2: Hawks data file (some lines: the remaining lines are in the same format)

```
# A tibble: 3 x 2
  species species_name
  <chr>   <chr>
1 SS      Sharp-shinned
2 CH      Cooper
3 RT      Red-tailed
```

Figure 3: Dataframe `hawk_names`. Note that this has a third species that might have been observed but was not.

```
t.test(weight ~ species, data = hawks)
```

```

    Welch Two Sample t-test

data:  weight by species
t = -32.215, df = 93.635, p-value < 2.2e-16
alternative hypothesis: true difference in means between group CH and group RT is not equal
95 percent confidence interval:
 -715.4837 -632.4050
sample estimates:
mean in group CH mean in group RT
        420.4857        1094.4301
```

Figure 4: Hawks data test

```
# A tibble: 22 x 2
   expend stature
    <dbl> <fct>
 1   9.21 obese
 2   7.53 lean
 3   7.48 lean
 4   8.08 lean
 5   8.09 lean
 6  10.2  lean
 7   8.4  lean
 8  10.9  lean
 9   6.13 lean
10   7.9  lean
11  11.5  obese
12  12.8  obese
13   7.05 lean
14  11.8  obese
15   9.97 obese
16   7.48 lean
17   8.79 obese
18   9.69 obese
19   9.68 obese
20   7.58 lean
21   9.19 obese
22   8.11 lean
```

Figure 5: Energy usage data (all)

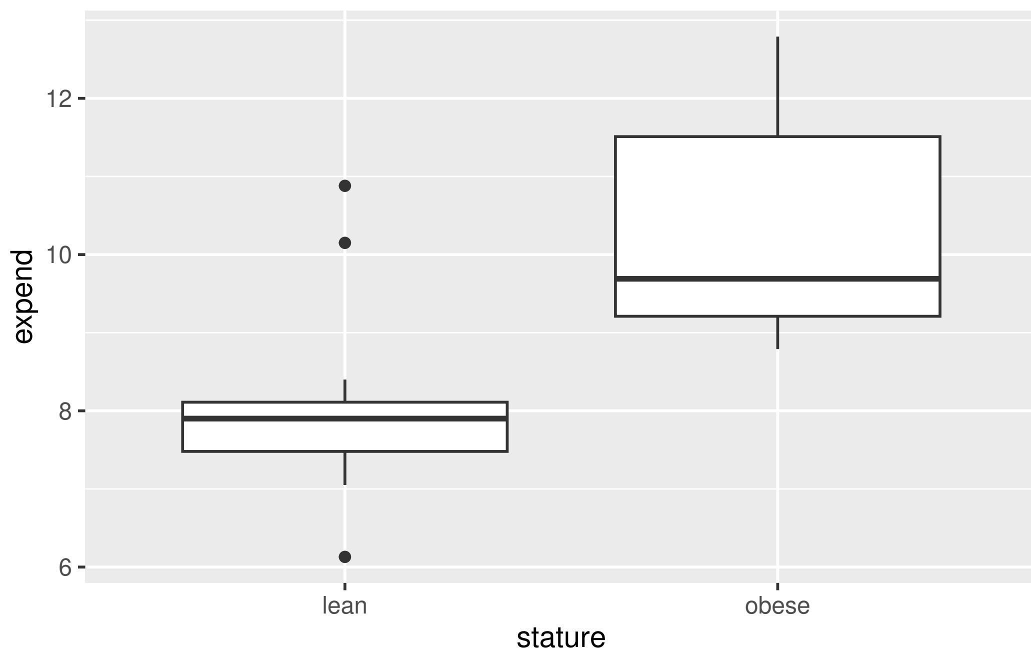

Figure 6: Energy usage graph

```
$grand_median
[1] 8.595

$table
       above
group   above below
  lean      2    11
  obese     9     0

$test
       what        value
1 statistic 1.523077e+01
2        df 1.000000e+00
3   P-value 9.514059e-05
```

Figure 7: Energy usage test

```
# A tibble: 20 x 4
   Subject Educ  Income2005 tr_income
     <int> <chr>      <int>     <dbl>
 1    2188 13-15      36000     13.8
 2    1436 012        10000     10
 3    5307 12         28000     12.9
 4    1222 12         38000     14.0
 5    2944 12         32000     13.4
 6    2809 16+        15000     11.1
 7    1778 13-15       8000      9.46
 8    3036 12         52000     15.1
 9    3023 12         30000     13.2
10    1148 012        31000     13.3
11   12016 16        150000     19.7
12    3949 13-15      39000     14.1
13    3850 12         32000     13.4
14    2057 13-15       9000      9.74
15    3609 13-15      39000     14.1
16    2333 13-15      85000     17.1
17     612 16         46136     14.7
18    3090 12         31600     13.3
19    4822 12         30775     13.2
20    2852 12         33000     13.5
```

Figure 8: Income and education data (20 randomly chosen rows)

```
# A tibble: 5 x 4
  Educ      n income_mean income_sd
  <chr> <int>       <dbl>     <dbl>
1 012     136        12.2      2.66
2 12     1020        13.2      2.55
3 13-15   648        13.8      2.84
4 16      406        15.3      3.30
5 16+     374        15.7      3.35
```

Figure 9: Summary data of transformed income for each education category

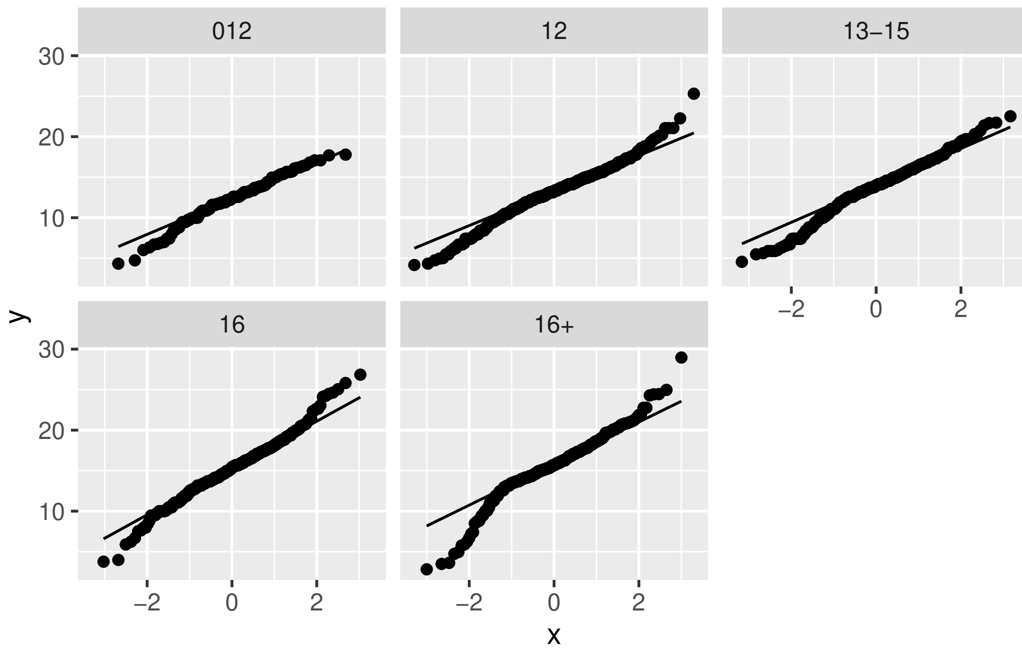

Figure 10: Plots of transformed income

```
              Df Sum Sq Mean Sq F value Pr(>F)
Educ           4   2902   725.4   87.44 <2e-16 ***
Residuals   2579  21395     8.3
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = tr_income ~ Educ, data = incomes)

$Educ
               diff        lwr       upr     p adj
12-012    0.9530098  0.2352841 1.6707356 0.0027287
13-15-012 1.5528225  0.8112558 2.2943892 0.0000001
16-012    3.0464410  2.2674789 3.8254031 0.0000000
16+-012   3.4879841  2.7007045 4.2752637 0.0000000
13-15-12  0.5998126  0.2048468 0.9947785 0.0003375
16-12     2.0934312  1.6320649 2.5547974 0.0000000
16+-12    2.5349742  2.0596996 3.0102489 0.0000000
16-13-15  1.4936185  0.9959743 1.9912627 0.0000000
16+-13-15 1.9351616  1.4245963 2.4457269 0.0000000
16+-16    0.4415431 -0.1219617 1.0050479 0.2039186
```

Figure 11: Income data, analysis 1

```
    One-way analysis of means (not assuming equal variances)

data:  tr_income and Educ
F = 77.507, num df = 4.00, denom df = 677.54, p-value < 2.2e-16


    Pairwise comparisons using Games-Howell test

data: tr_income by factor(Educ)

      012     12      13-15   16
12    0.00111 -       -       -
13-15 4.9e-08 0.00013 -       -
16    7.2e-13 3.9e-10 < 2e-16 -
16+   8.9e-13 3.6e-10 < 2e-16 0.34356


P value adjustment method: none

alternative hypothesis: two.sided
```

Figure 12: Income data, analysis 2

```
# A tibble: 3 x 3
      a     b     r
  <dbl> <dbl> <dbl>
1    12    14     1
2    13    15     2
3    16    17     3
```

```
d1 %>% pivot_longer(a:b, names_to = "trt", values_to = "yield")
```

Figure 13: A dataframe d1 and some code

```
# A tibble: 2 x 4
    x_1   x_2   y_1   y_2
  <dbl> <dbl> <dbl> <dbl>
1    21    22    11    12
2    23    24    13    14
```

Figure 14: Dataframe d2

```
# A tibble: 4 x 3
  time      x     y
  <chr> <dbl> <dbl>
1 1        21    11
2 2        22    12
3 1        23    13
4 2        24    14
```

Figure 15: Dataframe d3

```
# A tibble: 6 x 3
  rep   g         x
  <chr> <chr> <dbl>
1 r1    trt1     10
2 r2    trt1     11
3 r3    trt1     12
4 r1    trt2     13
5 r2    trt2     14
6 r3    trt2     15
```

Figure 16: Dataframe d4

```
# A tibble: 3 x 3
  rep    trt1  trt2
  <chr> <dbl> <dbl>
1 r1       10    13
2 r2       11    14
3 r3       12    15
```

Figure 17: Dataframe d5

```
# A tibble: 4 x 3
  u     v         w
  <chr> <chr> <dbl>
1 a     g1       10
2 b     g2       11
3 c     g2       13
4 b     g1       14
```

```
d6 %>% pivot_wider(names_from = v, values_from = w)
```

Figure 18: Dataframe `d6` and some code using `d6`

```
# A tibble: 14 x 2
   body_weight   mcr
         <int> <int>
 1         110   235
 2         110   198
 3         110   173
 4         230   174
 5         230   149
 6         230   124
 7         360   115
 8         360   130
 9         360   102
10         360    95
11         505   122
12         505   112
13         505    98
14         505    96
```

Figure 19: Cattle growth data

```
ggplot(growth, aes(x = body_weight, y = mcr)) + geom_point()
```

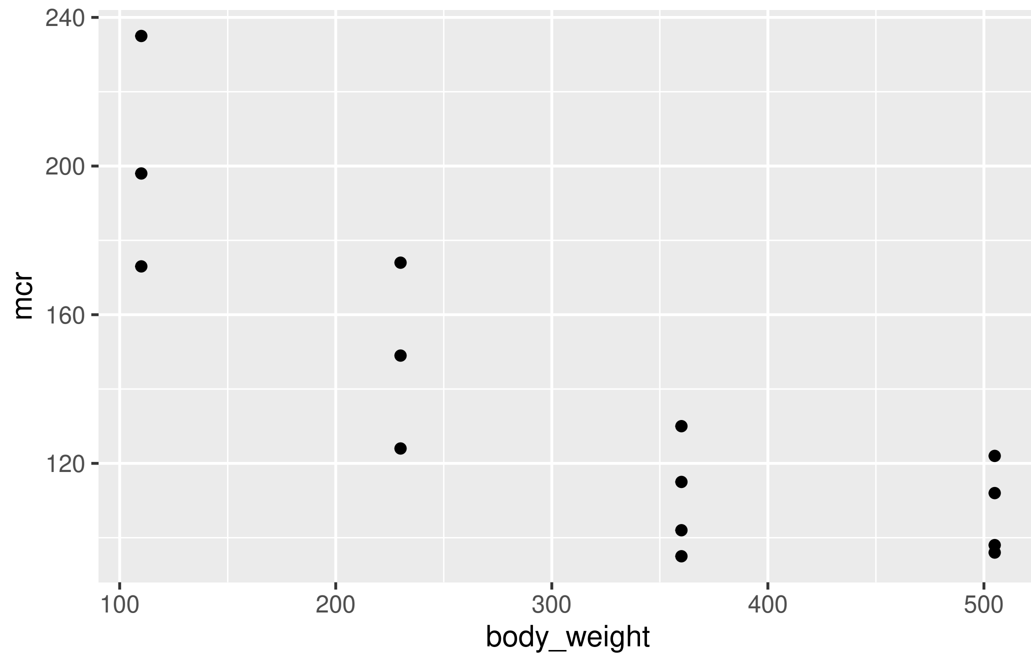

Figure 20: Cattle growth scatterplot

```r
growth.1 <- lm(mcr ~ body_weight, data = growth)
summary(growth.1)
```

```

Call:
lm(formula = mcr ~ body_weight, data = growth)

Residuals:
    Min      1Q  Median      3Q     Max 
-34.553 -13.595   2.138  14.381  48.185 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
(Intercept) 212.72093   15.78406   13.48 1.31e-08 ***
body_weight  -0.23551    0.04486   -5.25 0.000204 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 24.56 on 12 degrees of freedom
Multiple R-squared:  0.6967,	Adjusted R-squared:  0.6714 
F-statistic: 27.57 on 1 and 12 DF,  p-value: 0.0002043
```

Figure 21: Cattle growth regression analysis 1

```
ggplot(growth.1, aes(x = .fitted, y = .resid)) + geom_point()
```

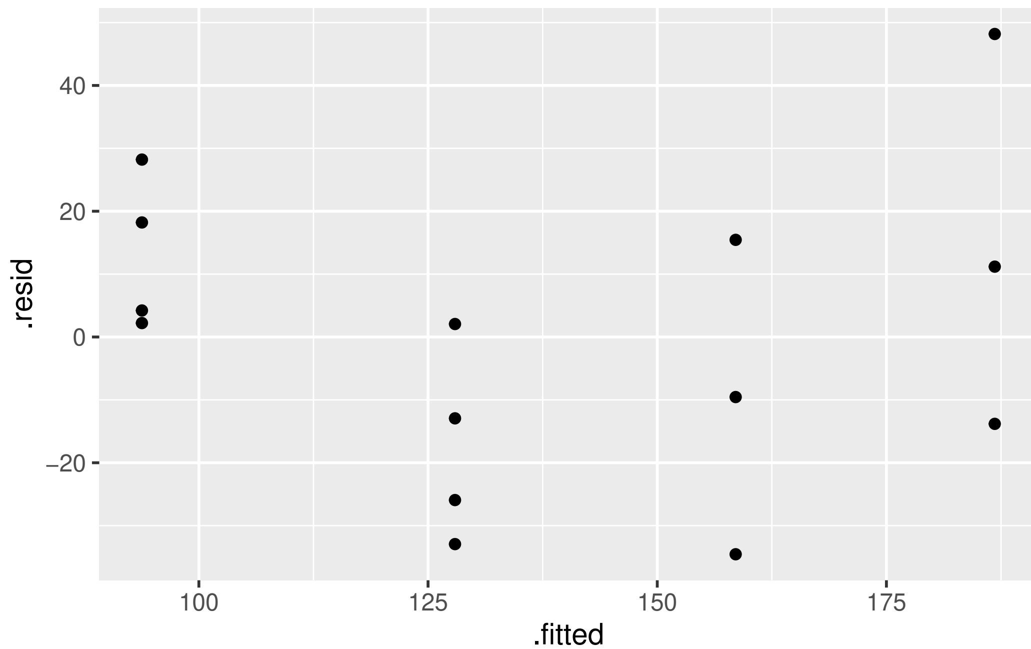

Figure 22: Cattle growth regression plot 1

```r
growth.2 <- lm(mcr ~ body_weight + I(body_weight^2), data = growth)
summary(growth.2)
```

```

Call:
lm(formula = mcr ~ body_weight + I(body_weight^2), data = growth)

Residuals:
   Min     1Q Median     3Q    Max 
-29.89 -10.42  -1.22  13.00  32.12 

Coefficients:
                   Estimate Std. Error t value Pr(>|t|)    
(Intercept)       2.753e+02  2.671e+01  10.305 5.47e-07 ***
body_weight      -7.481e-01  1.954e-01  -3.828   0.0028 ** 
I(body_weight^2)  8.197e-04  3.070e-04   2.670   0.0218 *  
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 19.99 on 11 degrees of freedom
Multiple R-squared:  0.816,	Adjusted R-squared:  0.7825 
F-statistic: 24.39 on 2 and 11 DF,  p-value: 9.051e-05
```

Figure 23: Cattle growth regression analysis

treatment	body	gut_length	mouthpart_damage
Bd	20.28	191.40	0.679
Bd	19.75	142.92	0.607
Bd	19.28	169.81	0.750
Bd	17.81	152.18	0.821
Bd	20.79	171.33	0.696
Bd	16.99	153.92	0.571
Bd	18.12	181.08	0.571
Bd	16.95	124.90	0.571
Bd	19.45	173.06	0.679
Bd	19.44	207.01	0.786
Bd	18.32	143.77	0.750
Bd	20.27	220.35	0.714
Bd	18.71	130.00	0.714
Bd	19.05	195.13	0.714
Control	19.46	177.92	0.421
Control	22.62	181.58	0.546
Control	19.42	154.29	0.414
Control	19.55	217.67	0.714
Control	19.69	185.88	0.643
Control	22.40	249.29	0.714
Control	20.78	196.90	0.714
Control	18.78	202.16	0.693
Control	22.48	210.86	0.554
Control	23.18	215.60	0.857
Control	18.59	174.16	0.607
Control	25.92	222.83	0.643
Control	20.38	228.74	0.643

Figure 24: Tadpole data

```
tadpoles.1 <- lm(gut_length ~ body + mouthpart_damage + treatment,
                 data = tadpoles)
```

Figure 25: Tadpole regression

	Df	Sum of Sq	RSS	AIC	F value	Pr(>F)
	NA	NA	11829.09	172.2270	NA	NA
body	1	2830.183	14659.28	176.0188	5.502892	0.0279750
mouthpart_damage	1	2295.347	14124.44	175.0153	4.462978	0.0457019
treatment	1	2658.389	14487.48	175.7005	5.168863	0.0326488

Figure 26: Tadpole `drop1` output

```r
summary(tadpoles.1)
```

```

Call:
lm(formula = gut_length ~ body + mouthpart_damage + treatment,
    data = tadpoles)

Residuals:
    Min      1Q  Median      3Q     Max 
-39.422 -17.701  -6.771  16.338  40.877 

Coefficients:
                 Estimate Std. Error t value Pr(>|t|)  
(Intercept)       -20.258     53.070  -0.382   0.7062  
body                6.442      2.746   2.346   0.0280 *
mouthpart_damage   96.839     45.839   2.113   0.0457 *
treatmentControl   25.412     11.177   2.274   0.0326 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 22.68 on 23 degrees of freedom
Multiple R-squared:  0.5502,	Adjusted R-squared:  0.4915 
F-statistic: 9.378 on 3 and 23 DF,  p-value: 0.0003092
```

Figure 27: Tadpole `summary` output

```
ggplot(tadpoles.1, aes(x = .fitted, y = .resid)) + geom_point()
```

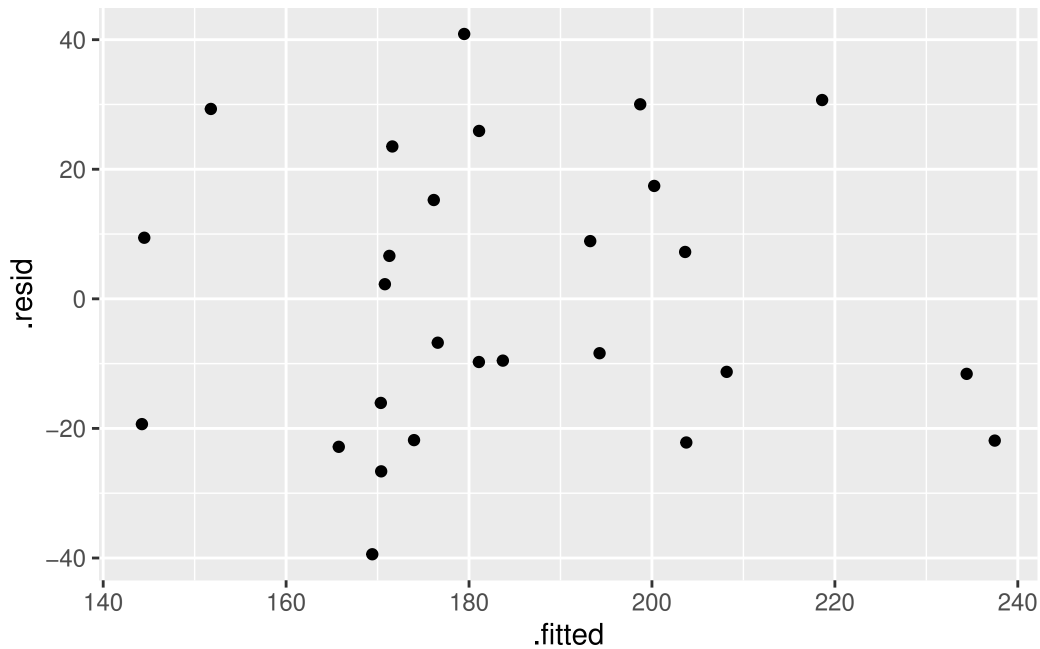

Figure 28: Tadpoles plot 1

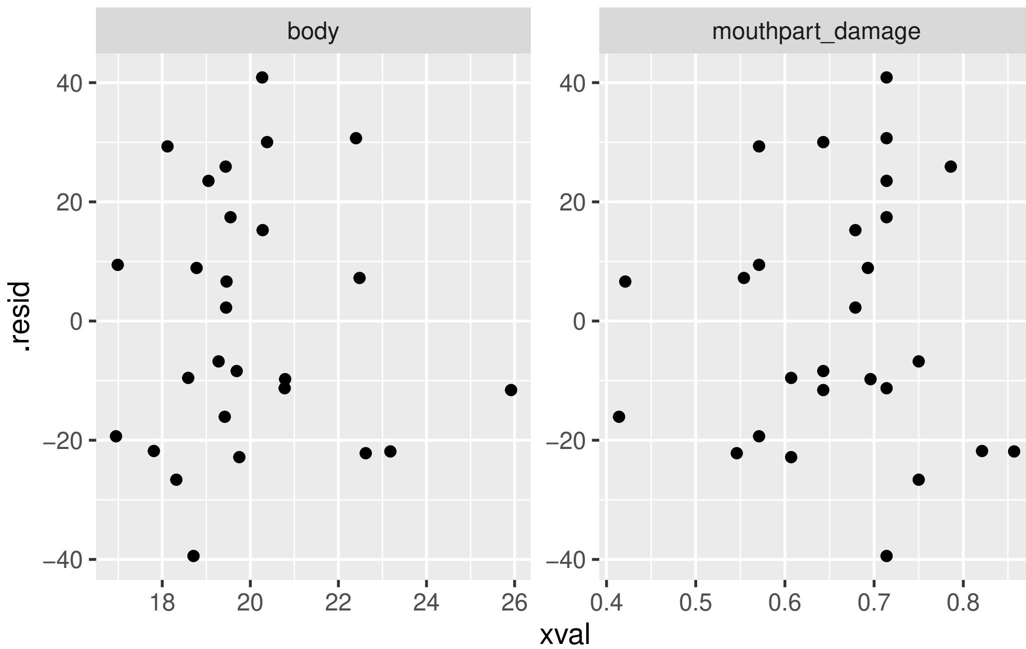

Figure 29: Tadpoles plot 2

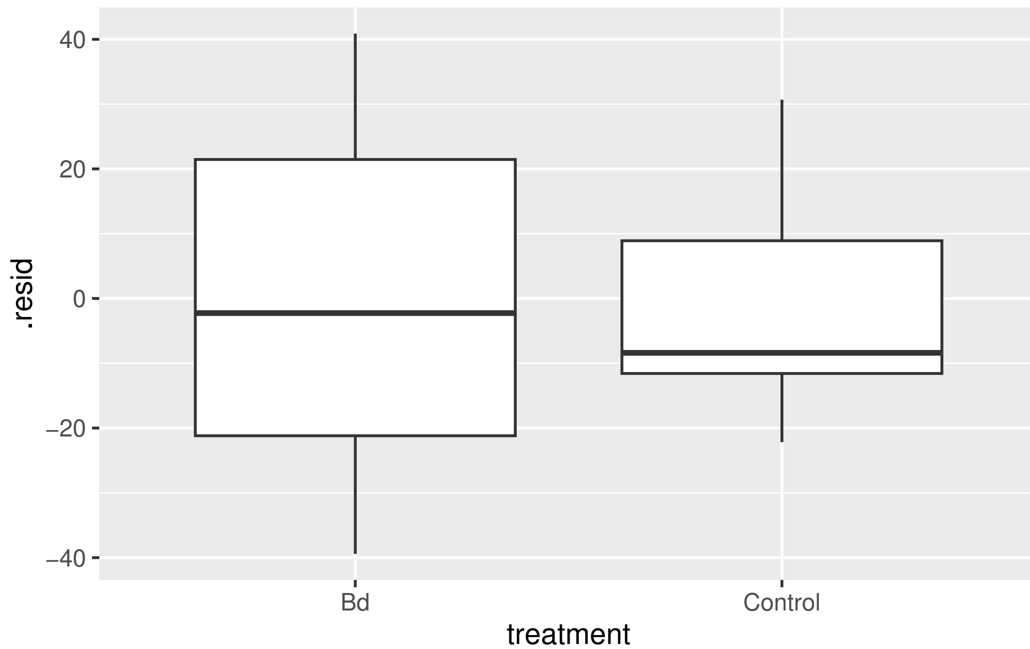

Figure 30: Tadpoles plot 3