Wasp Reproduction Experiment

May 27, 2022

Stella Aurelia

Agenda

- ❖ Introduction to the Study
- ❖ Project Objective
- ❖ Understanding the Experiment
  - ➢ Methodology
- ❖ Descriptive Analysis
- ❖ Proposed Statistical Method
  - ➢ Data analysis
  - ➢ Post-hoc Test
- ❖ Conclusion

Introduction To the Study: Controlling Pest Population

When we use natural enemies to reduce invasive species population, we refer the natural enemies as "biological control agents".

Female wasps (biocontrol agent) lay their eggs inside pests eggs and the developing parasitoid larvae kill pest eggs by feeding inside the pest egg.

Project objective:
Test if some plants can increase reproduction of wasps in order to control the pest population.

Understanding the Experiment

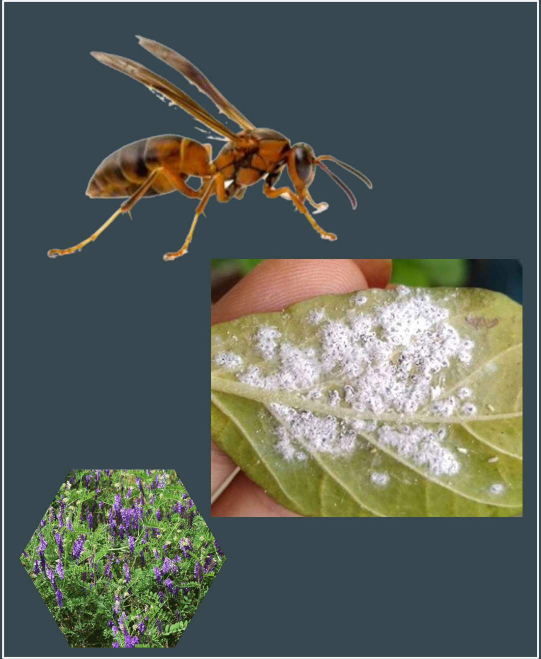
3 Treatments - Can plants increase reproduction of wasps?

Buckwheat (BW)

No Plant/ Water Only (W)

Vetch (V)

Methodology



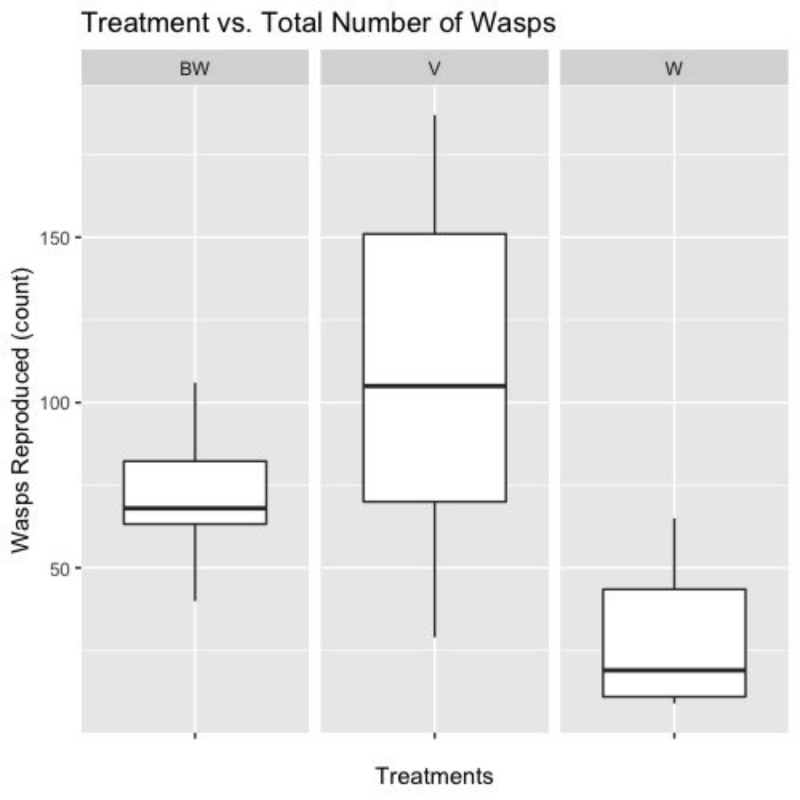
Descriptive Analysis

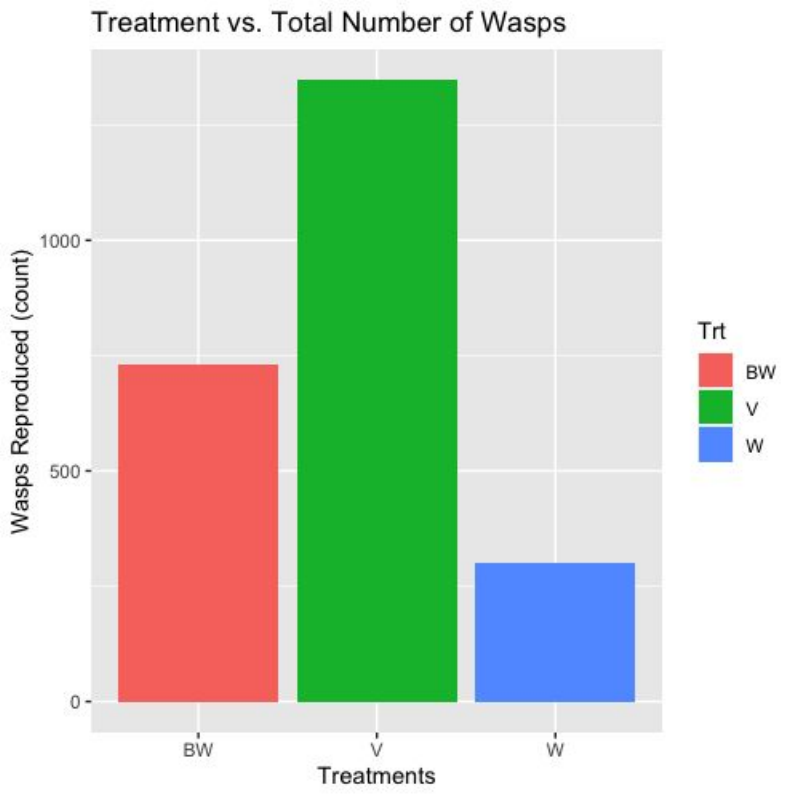
Trt	Rep	Longevity	Male	Female	Total	Sex ratio
BW	1	10	15	38	53	71.69811
BW	3	14	22	41	63	65.07937
BW	7	12	30	47	77	61.03896
BW	9	14	27	13	40	32.5
BW	10	14	54	10	64	15.625
BW	12	13	105	1	106	0.943396
BW	14	17	.	.	68	.
BW	15	14	39	29	68	42.64706
BW	16	15	50	56	106	52.83019
BW	17	14	35	49	84	58.33333
W	1	8	11	6	17	35.29412
W	2	14	5	14	19	73.68421
W	4	3	4	5	9	55.55556
W	5	14	27	19	46	41.30435
W	6	9	38	3	41	7.317073
W	1	12	38	13	51	25.4902
W	2	11	9	1	10	10
W	3	17	10	1	11	9.090909
W	4	5	8	12	20	60
W	7	10	39	26	65	40
W	8	16	9	2	11	18.18182
V	2	14	78	10	88	11.36364
V	4	16	22	7	29	24.13793
V	5	14	23	14	37	37.83784
V	7	16	92	32	124	25.80645
V	8	17	98	53	151	35.09934
V	9	14	55	50	105	47.61905
V	10	16	95	59	154	38.31169
V	12	17	13	20	33	60.60606
V	13	12	81	92	173	53.17919
V	14	18	173	14	187	7.486631
V	16	16	101	18	119	15.12605
V	17	14	68	2	70	2.857143
V	18	15	44	33	77	42.85714

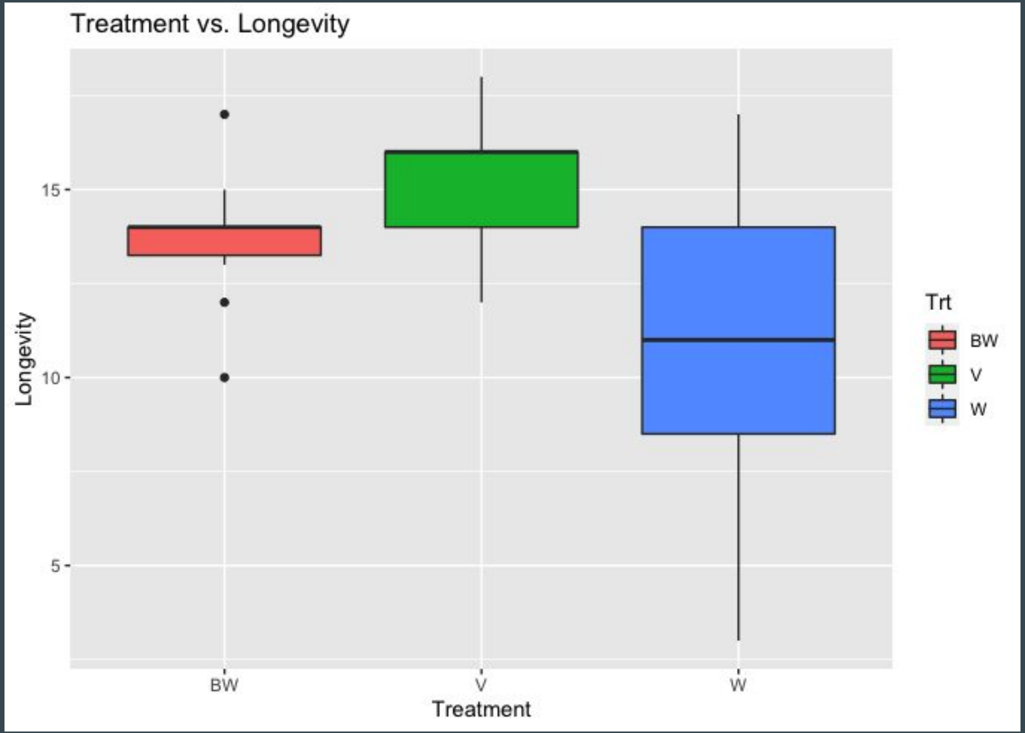
Independent Variable:
-Trt

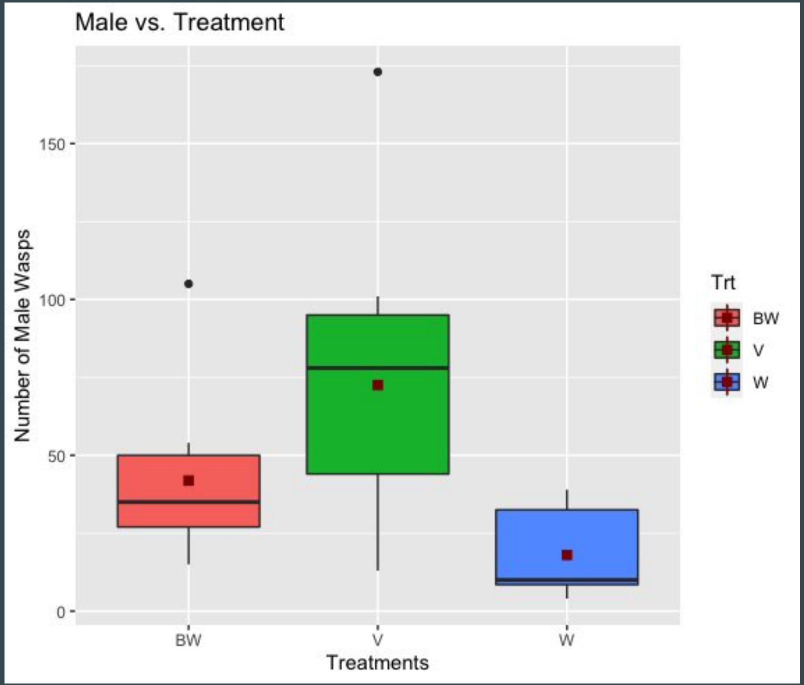
Dependent Variable:
-Longevity
-Male
-Female
-Total
-Sex ratio

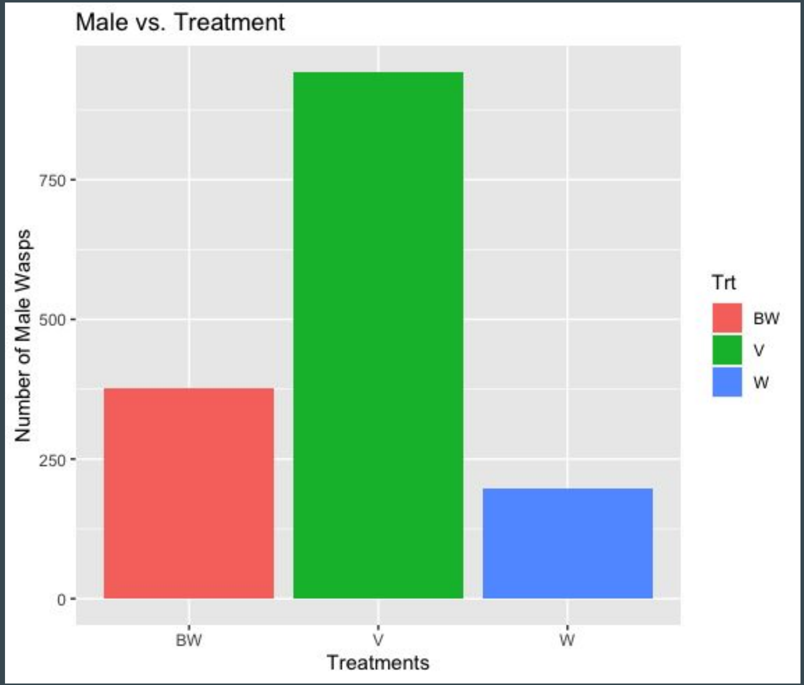
Ignore:
-Rep

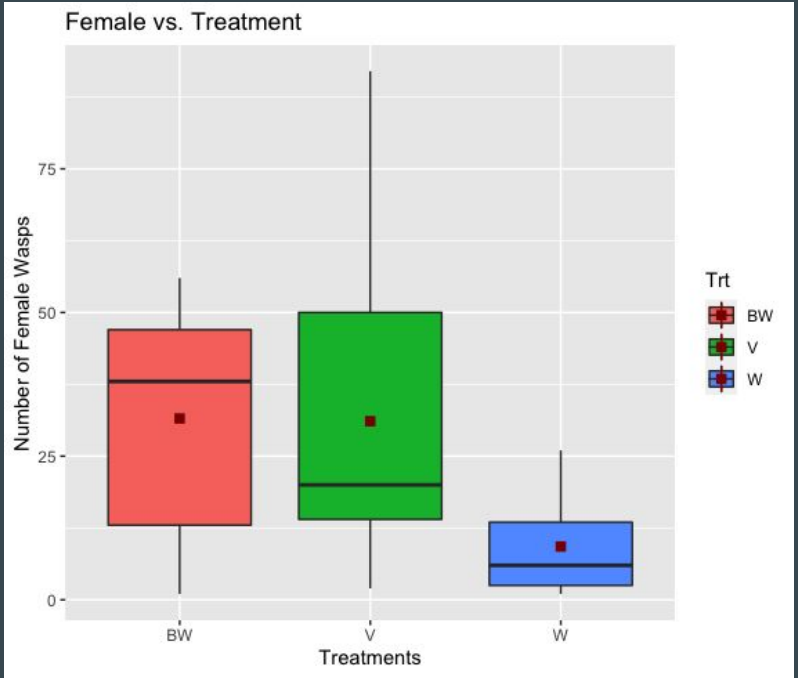

Treatment vs. Total Number of Wasps
BW
V
W
150
100
50
Wasps Reproduced (count)
Treatments


Treatment vs. Total Number of Wasps
1000
500
0
Wasps Reproduced (count)
BW
V
W
Treatments
Trt
BW
V
W

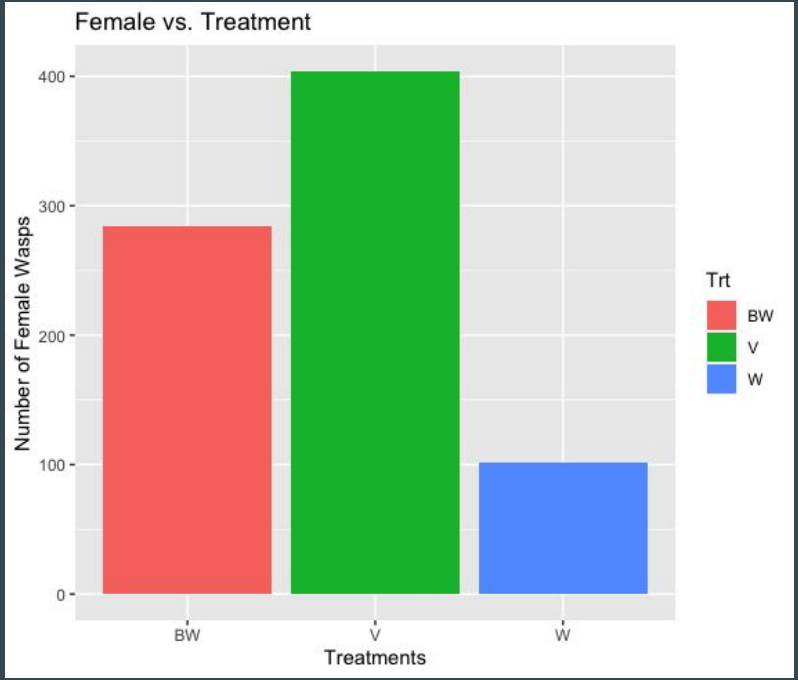

Treatment vs. Longevity
15
10
5
Longevity
BW
V
W
Treatment
Trt
BW
V
W


Male vs. Treatment
Number of Male Wasps
150
100
50
0
BW
V
W
Treatments
Trt
BW
V
W


Male vs. Treatment
Number of Male Wasps
750
500
250
0
BW
V
W
Treatments
Trt
BW
V
W

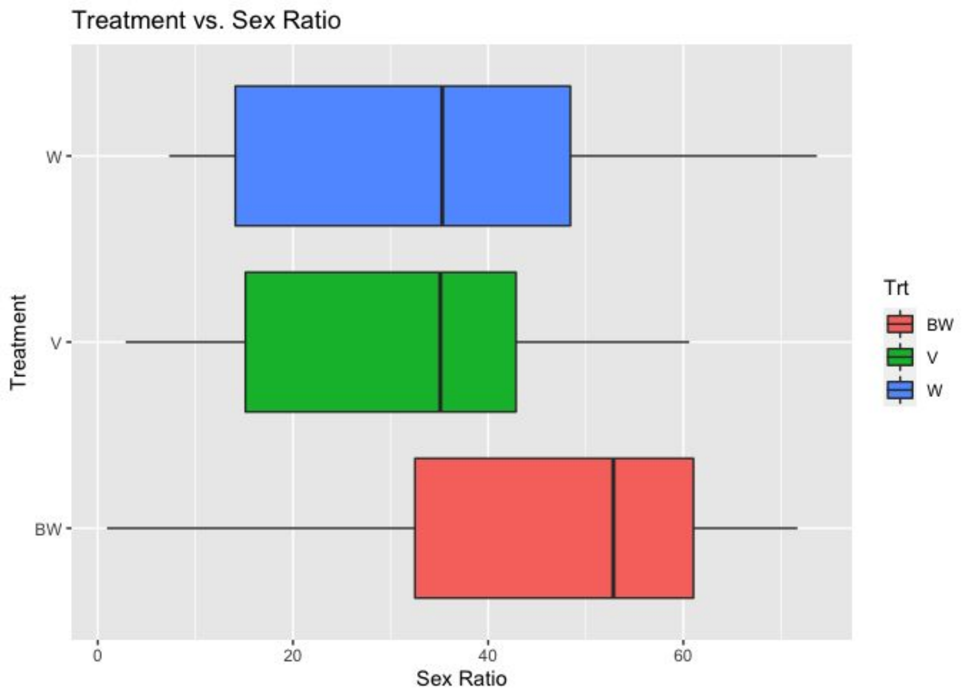


Female vs. Treatment
Number of Female Wasps
0
25
50
75
BW
V
W
Treatments
Trt
BW
V
W



Female vs. Treatment
Number of Female Wasps
0
100
200
300
400
BW
V
W
Treatments
Trt
BW
V
W

Sex Ratio = # of female/ # total



Proposed Statistical Method

Anova Model

Studying the effects of the three different treatments on each variable.

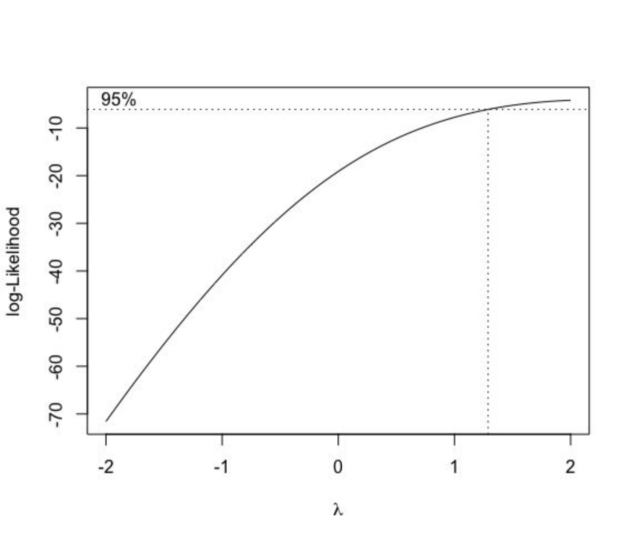
Assumptions:
- Samples are independent
- Normal Distribution
- Equal Variance

H_0: $\mu_{BW} = \mu_W = \mu_V$

H_a: at least one is different

Data Analysis for Each Variable

Transformation: Longevity Variable




```
> lambda <- bc$x[which.max(bc$y)]
> lambda
[1] 2
```

Lambda value is 2 so use square-transformation on the Y variable.

Data Analysis: Longevity Variable

```
> shapiro.test(Mealy_Data$Longevity^2)

        Shapiro-Wilk normality test

data:  Mealy_Data$Longevity^2
W = 0.9455, p-value = 0.09001
```

```
> bartlett.test(Longevity^2 ~ Trt, data = Mealy_Data)

        Bartlett test of homogeneity of variances

data:  Longevity^2 by Trt
Bartlett's K-squared = 5.2402, df = 2, p-value = 0.07279
```

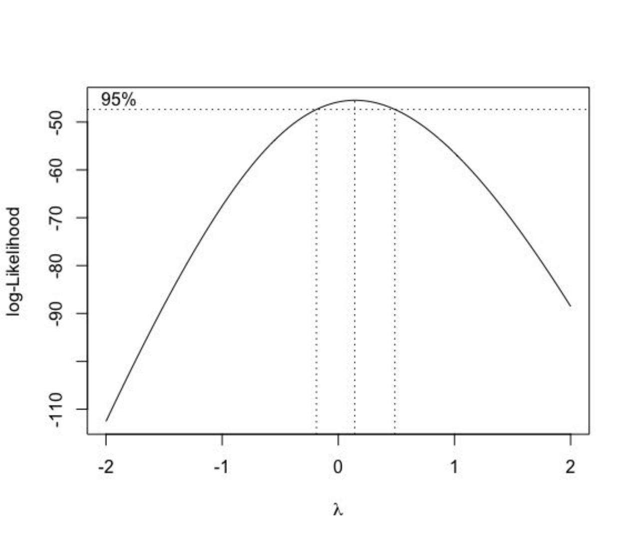
Assumptions are met.

Data Analysis: Longevity Variable

```
> summary(aov(Longevity^2 ~ Trt, data = Mealy_Data))
            Df Sum Sq Mean Sq F value Pr(>F)
Trt          2  62249   31125   7.146 0.0028 **
Residuals   31 135028    4356
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Since the p-value < $\alpha = 0.05$, we reject the null hypothesis. We can conclude that there is a significant difference in treatment effects on the longevity variable.

Transformation: Male Variable




```
> lambda2 <- bc2$x[which.max(bc2$y)]
> lambda2
[1] 0.1414141
```

Lambda value is 0.1414141 so use log transformation on the Y variable.

Data Analysis: Male Variable

```
> shapiro.test(log(as.numeric(Mealy_Data$Male)))

        Shapiro-Wilk normality test

data:  log(as.numeric(Mealy_Data$Male))
W = 0.96595, p-value = 0.3772
```

```
> bartlett.test(log(as.numeric(Mealy_Data$Male)) ~ Trt, data = Mealy_Data)

        Bartlett test of homogeneity of variances

data:  log(as.numeric(Mealy_Data$Male)) by Trt
Bartlett's K-squared = 1.2273, df = 2, p-value = 0.5414
```

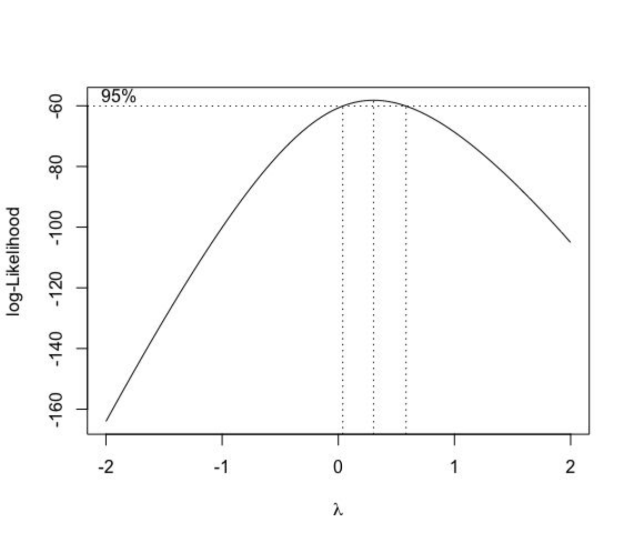
Assumptions are met.

Data Analysis: Male Variable

```
> summary(aov(log(as.numeric(Mealy_Data$Male)) ~ Trt, data = Mealy_Data))
            Df Sum Sq Mean Sq F value   Pr(>F)    
Trt          2  13.54   6.768   12.64 0.000104 ***
Residuals   30  16.06   0.535                     
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
1 observation deleted due to missingness
```

Since the p-value < α = 0.05, we reject the null hypothesis. We can conclude that there is a significant difference in treatment effects on the male variable.

Transformation: Female Variable




```
> lambda3 <- bc3$x[which.max(bc3$y)]
> lambda3
[1] 0.3030303
```

Lambda value is 0.3030303 so use log transformation on the Y variable. But since log-transformation still violates assumption, use square-root transformation instead.

Data Analysis: Female Variable

```
> Mealy_Data$Female2 <- sqrt(Mealy_Data$Female)
> shapiro.test(Mealy_Data$Female2)

        Shapiro-Wilk normality test

data:  Mealy_Data$Female2
W = 0.96047, p-value = 0.2666
```

```
> bartlett.test(Female2 ~ Trt, data = Mealy_Data)

        Bartlett test of homogeneity of variances

data:  Female2 by Trt
Bartlett's K-squared = 2.5973, df = 2, p-value = 0.2729
```

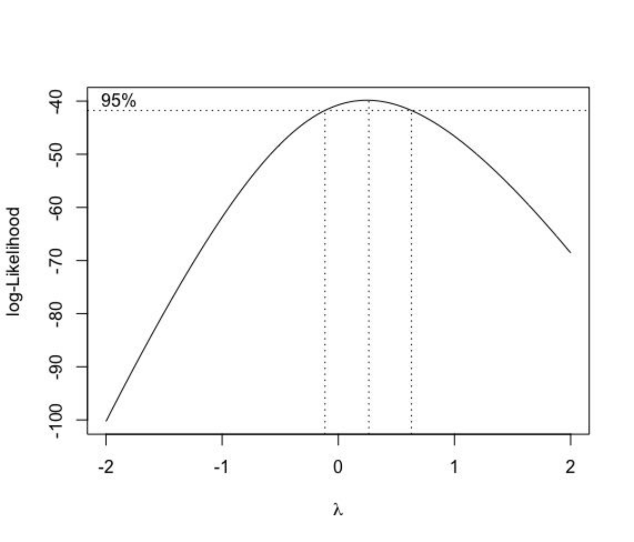
Assumptions are met.

Data Analysis: Female Variable

```
> summary(aov(Female2 ~ Trt, data = Mealy_Data))
            Df Sum Sq Mean Sq F value Pr(>F)
Trt          2  43.12  21.561   5.304 0.0107 *
Residuals   30 121.95   4.065
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Since the p-value < α = 0.05, we reject the null hypothesis. We can conclude that there is a significant difference in treatment effects on the female variable.

Transformation: Total Variable




```
> lambda4 <- bc4$x[which.max(bc4$y)]
> lambda4
[1] 0.2626263
```

Lambda value is 0.2626263 so use log transformation on the Y variable.

Data Analysis: Total Variable

```
> shapiro.test(log(Mealy_Data$Total))

	Shapiro-Wilk normality test

data:  log(Mealy_Data$Total)
W = 0.93694, p-value = 0.05001
```

```
> leveneTest(log(Total) ~ Trt, data = Mealy_Data)
Levene's Test for Homogeneity of Variance (center = median)
      Df F value Pr(>F)
group  2  3.1657 0.0561 .
      31
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Assumptions are met.

Data Analysis: Total Variable

```
> summary(aov(log(Total) ~ Trt, data = Mealy_Data))
            Df Sum Sq Mean Sq F value   Pr(>F)    
Trt          2  13.16   6.580    18.6 4.93e-06 ***
Residuals   31  10.97   0.354                     
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Since the p-value < α = 0.05, we reject the null hypothesis. We can conclude that there is a significant difference in treatment effects on the total variable.

Transformation: Sex Ratio Variable

No transformation needed.

Data Analysis: Sex Ratio Variable

```
> shapiro.test(Mealy_Data$`Sex ratio`)

        Shapiro-Wilk normality test

data:  Mealy_Data$`Sex ratio`
W = 0.9556, p-value = 0.1935
```

```
> bartlett.test(`Sex ratio` ~ Trt, data = Mealy_Data)

        Bartlett test of homogeneity of variances

data:  Sex ratio by Trt
Bartlett's K-squared = 0.8015, df = 2, p-value = 0.6698
```

Assumptions are met.

Data Analysis: Sex Ratio Variable

```
> summary(aov(`Sex ratio` ~ Trt, data = Mealy_Data))
            Df Sum Sq Mean Sq F value Pr(>F)
Trt          2   1020   509.9   1.127  0.337
Residuals   30  13572   452.4
```

Since the p-value $< \alpha = 0.05$, we reject the null hypothesis. We can conclude that there is a significant difference in treatment effects on the sex ratio variable.

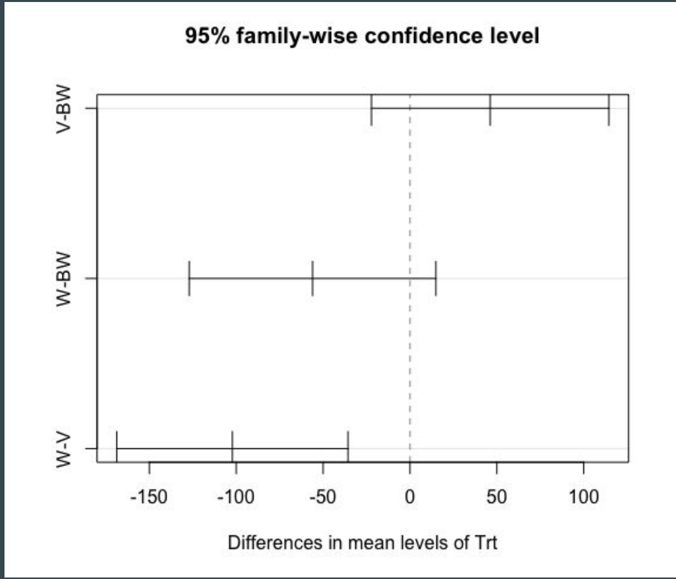
Post-Hoc Test: Multiple Comparisons Test

Longevity Variable

```
> TukeyHSD(a1, conf.level = 0.95)
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = Longevity^2 ~ Trt, data = Mealy_Data)

$Trt
           diff        lwr       upr     p adj
V-BW   46.14615  -22.17714 114.46945 0.2356059
W-BW  -56.06364 -127.03610  14.90882 0.1433992
W-V  -102.20979 -168.75460 -35.66499 0.0018861
```



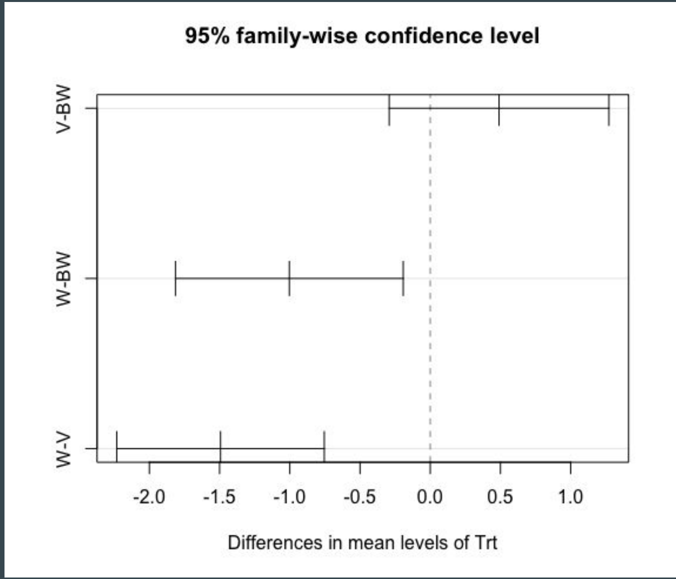
There is a significance between-group difference for treatments water and vetch.

Male Variable

```
> TukeyHSD(a2, conf.level = 0.95)
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = log(as.numeric(Mealy_Data$Male)) ~ Trt, data = Mealy_Data)

$Trt
           diff        lwr        upr     p adj
V-BW  0.4900378 -0.2921308  1.2722064 0.2850962
W-BW -1.0033129 -1.8140487 -0.1925771 0.0127842
W-V  -1.4933507 -2.2323084 -0.7543930 0.0000711
```



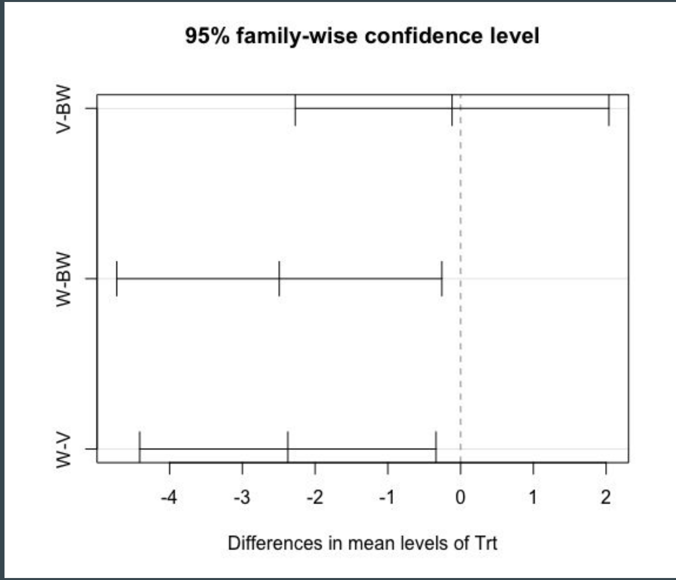
There is a significance between-group difference for treatments water and vetch as well as water and buckwheat.

Female Variable

```
> TukeyHSD(a3, conf.level = 0.95)
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = Female2 ~ Trt, data = Mealy_Data)

$Trt
           diff       lwr        upr     p adj
V-BW -0.1176229 -2.272909  2.0376633 0.9900737
W-BW -2.4923743 -4.726378 -0.2583704 0.0262540
W-V  -2.3747514 -4.410969 -0.3385340 0.0195643
```



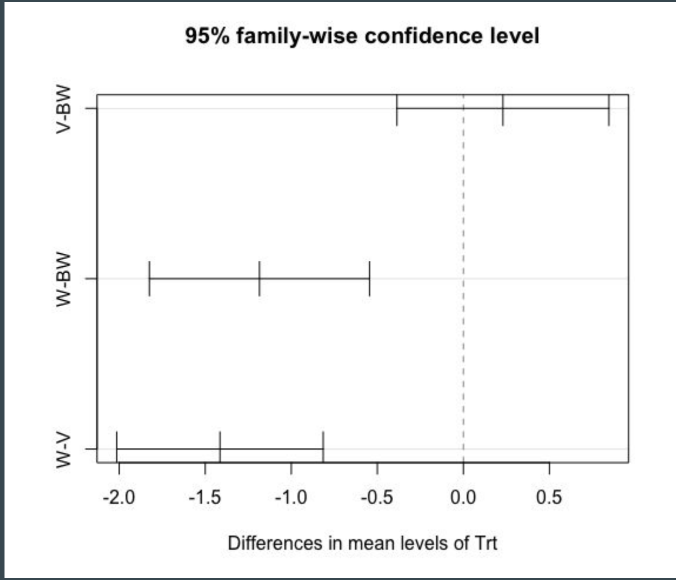
There is a significance between-group difference for treatments water and vetch as well as water and buckwheat.

Total Variable

```
> TukeyHSD(a4, conf.level = 0.95)
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = log(Total) ~ Trt, data = Mealy_Data)

$Trt
            diff        lwr        upr     p adj
V-BW   0.2295493 -0.386169  0.8452675 0.6334667
W-BW  -1.1849513 -1.824543 -0.5453592 0.0002177
W-V   -1.4145006 -2.014191 -0.8148097 0.0000063
```



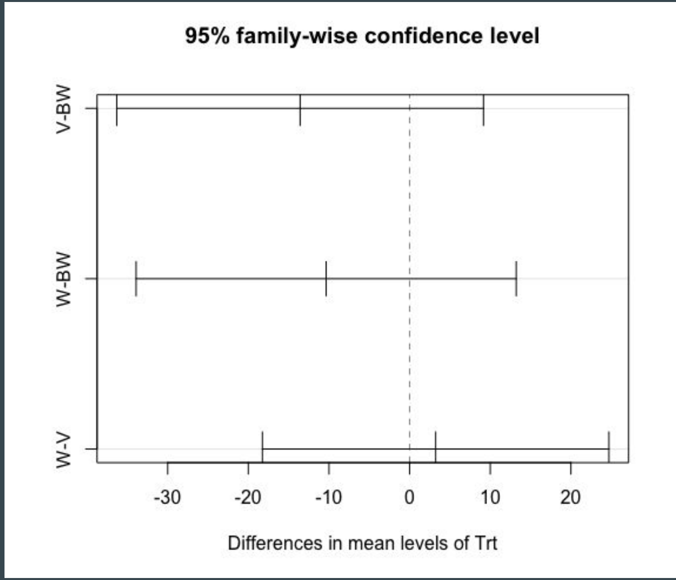
There is a significance between-group difference for treatments water and vetch as well as water and buckwheat.

Sex Ratio Variable

```
> TukeyHSD(sexratio.aov, conf.level=.95)
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = `Sex ratio` ~ Trt, data = Mealy_Data)

$Trt
           diff       lwr       upr     p adj
V-BW -13.576471 -36.31397  9.161024 0.3183623
W-BW -10.347329 -33.91527 13.220610 0.5321550
W-V    3.229142 -18.25222 24.710504 0.9272770
```



There is no significance between-group difference for any of the treatments.

Conclusion

There is a significant difference in treatment effects.

We can conclude that there is indeed an effect of plants increasing the reproduction of wasps to control the pest population.