

fruit flies

Chris Parrish

January 10, 2016

fruit flies

reference:

- Cannon, et al., Stat2, chapter 07, example 7.1, 7.4-7.9. 7.13-7.14 - Cannon, et al., Student's R Manual, chapter 07

Import the data.

```
data <- read.csv("FruitFlies.csv", header=TRUE)
head(data, 4)
```

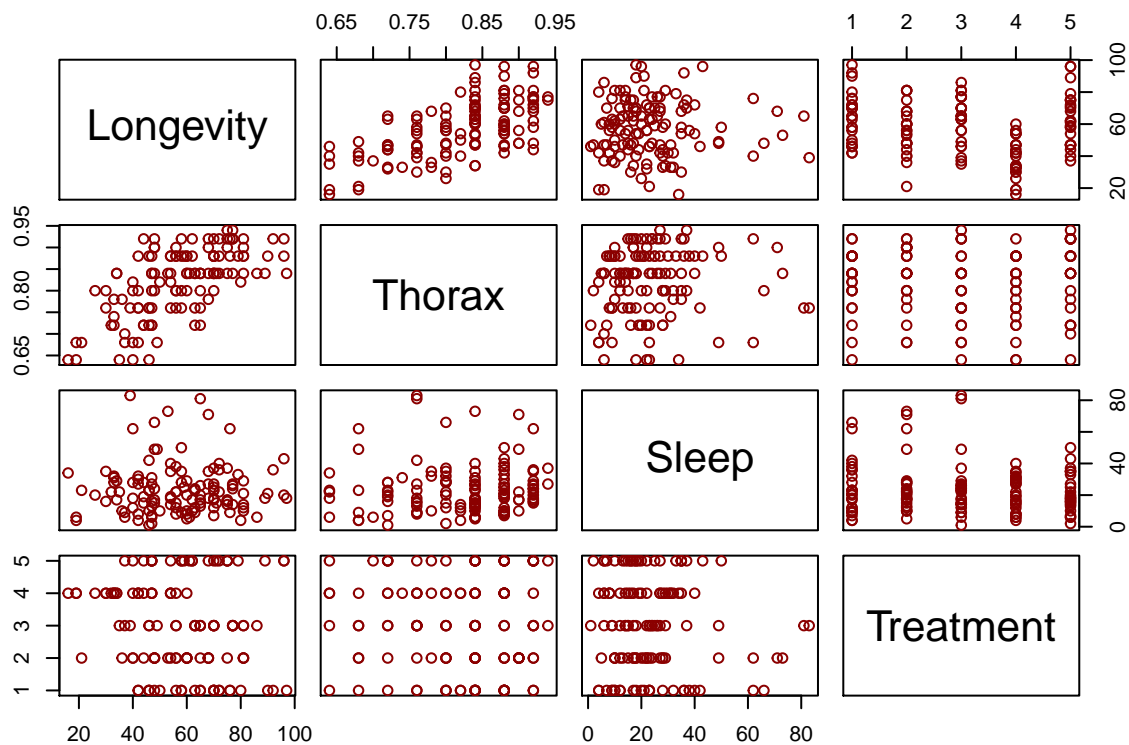
```
##   ID Partners Type Longevity Thorax Sleep Treatment
## 1  1         8   0        35  0.64  22 8 pregnant
## 2  2         8   0        37  0.68   9 8 pregnant
## 3  3         8   0        49  0.68  49 8 pregnant
## 4  4         8   0        46  0.72   1 8 pregnant
```

```
dim(data)
```

```
## [1] 125  7
```

Scatterplot matrix.

```
pairs(~ Longevity + Thorax + Sleep + Treatment, data=data, col="darkred")
```



Levene's test.

```
library(car)
leveneTest(Longevity ~ Treatment, data=data)

## Levene's Test for Homogeneity of Variance (center = median)
##      Df F value Pr(>F)
## group  4  0.4916 0.7419
##      120
```

Multiple comparisons: Fisher's LSD

```
with(data,
      pairwise.t.test(Longevity, Treatment, p.adj='none'))

##
## Pairwise comparisons using t tests with pooled SD
##
## data: Longevity and Treatment
##
##      1 pregnant 1 virgin 8 pregnant 8 virgin
## 1 virgin  0.057    -        -        -
## 8 pregnant 0.732    0.118    -        -
## 8 virgin  7.3e-09  3.4e-05  3.7e-08 -
## none     0.768    0.107    0.962    3.0e-08
##
## P value adjustment method: none
```

Multiple comparisons: Bonferroni

```
with(data,
      pairwise.t.test(Longevity, Treatment, p.adj='bonferroni'))

##
## Pairwise comparisons using t tests with pooled SD
##
## data: Longevity and Treatment
##
##      1 pregnant 1 virgin 8 pregnant 8 virgin
## 1 virgin  0.57282  -        -        -
## 8 pregnant 1.00000  1.00000  -        -
## 8 virgin  7.3e-08  0.00034  3.7e-07  -
## none     1.00000  1.00000  1.00000  3.0e-07
##
## P value adjustment method: bonferroni
```

Multiple comparisons: Tukey's HSD

```
fruitfly.aov <- aov(Longevity ~ Treatment, data=data)
options(show.signif.stars=FALSE)
summary(fruitfly.aov)
```

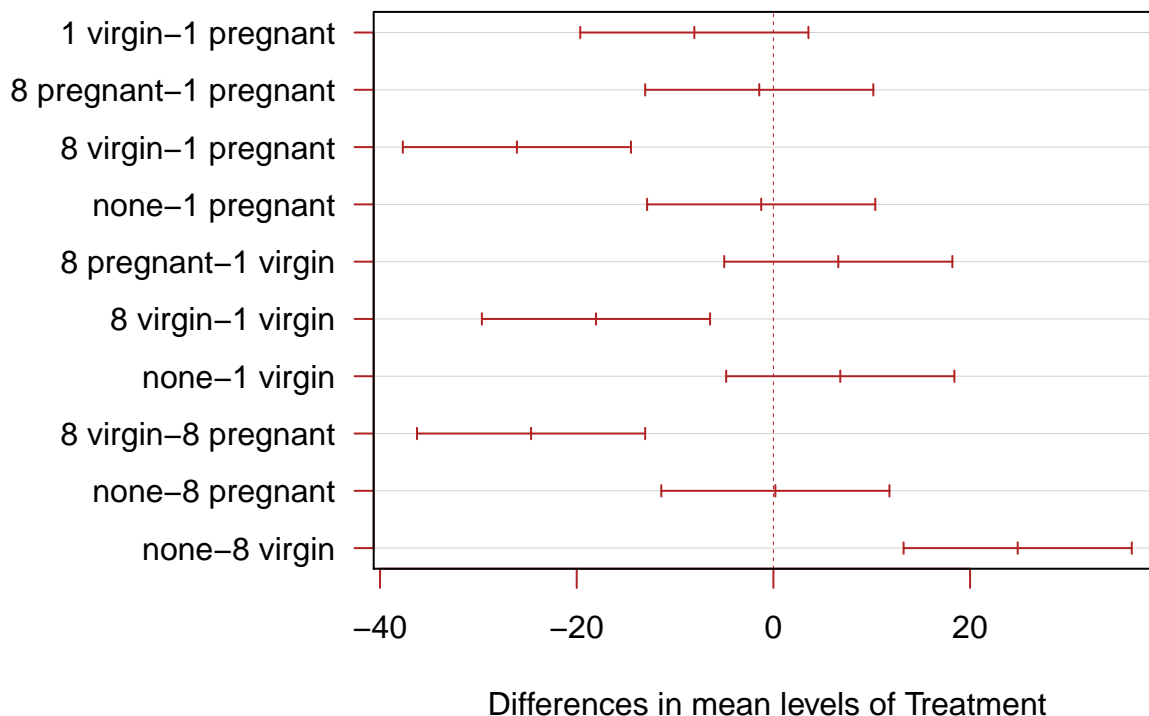
```
##           Df Sum Sq Mean Sq F value    Pr(>F)
## Treatment    4  11939   2984.8   13.61 3.52e-09
## Residuals  120   26314    219.3
```

```
TukeyHSD(fruitfly.aov)
```

```
## Tukey multiple comparisons of means
## 95% family-wise confidence level
##
## Fit: aov(formula = Longevity ~ Treatment, data = data)
##
## $Treatment
##           diff          lwr          upr      p adj
## 1 virgin-1 pregnant   -8.04 -19.640468   3.560468 0.3126549
## 8 pregnant-1 pregnant  -1.44 -13.040468  10.160468 0.9969591
## 8 virgin-1 pregnant  -26.08 -37.680468 -14.479532 0.0000001
## none-1 pregnant     -1.24 -12.840468  10.360468 0.9983034
## 8 pregnant-1 virgin    6.60  -5.000468  18.200468 0.5157692
## 8 virgin-1 virgin    -18.04 -29.640468  -6.439532 0.0003240
## none-1 virgin         6.80  -4.800468  18.400468 0.4854206
## 8 virgin-8 pregnant  -24.64 -36.240468 -13.039532 0.0000004
## none-8 pregnant       0.20 -11.400468  11.800468 0.9999988
## none-8 virgin        24.84  13.239532  36.440468 0.0000003
```

```
oldpar <- par(mar=c(4, 10, 4, 2)) # make more room for names
plot(TukeyHSD(fruitfly.aov), las=1, col="firebrick")
```

95% family-wise confidence level



```
par(oldpar)
```

Multiple comparisons: Fisher's LSD using mc.R

#	i	j	n.i	n.j	i-j.diff	lcl	ucl	signif?
# 1	1pg	1vir	25	25	8.04	-0.2526709	16.332671	no
# 2	1pg	8pg	25	25	1.44	-6.8526709	9.732671	no
# 3	1pg	8vir	25	25	26.08	17.7873291	34.372671	YES
# 4	1pg	none	25	25	1.24	-7.0526709	9.532671	no
# 5	1vir	8pg	25	25	-6.60	-14.8926709	1.692671	no
# 6	1vir	8vir	25	25	18.04	9.7473291	26.332671	YES
# 7	1vir	none	25	25	-6.80	-15.0926709	1.492671	no
# 8	8pg	8vir	25	25	24.64	16.3473291	32.932671	YES
# 9	8pg	none	25	25	-0.20	-8.4926709	8.092671	no
# 10	8vir	none	25	25	-24.84	-33.1326709	-16.547329	YES

Multiple comparisons: Bonferroni using mc.R

#	i	j	n.i	n.j	i-j.diff	lcl	ucl	signif?
# 1	1pg	1vir	25	25	8.04	-3.938157	20.018157	no
# 2	1pg	8pg	25	25	1.44	-10.538157	13.418157	no
# 3	1pg	8vir	25	25	26.08	14.101843	38.058157	YES
# 4	1pg	none	25	25	1.24	-10.738157	13.218157	no
# 5	1vir	8pg	25	25	-6.60	-18.578157	5.378157	no
# 6	1vir	8vir	25	25	18.04	6.061843	30.018157	YES
# 7	1vir	none	25	25	-6.80	-18.778157	5.178157	no
# 8	8pg	8vir	25	25	24.64	12.661843	36.618157	YES
# 9	8pg	none	25	25	-0.20	-12.178157	11.778157	no
# 10	8vir	none	25	25	-24.84	-36.818157	-12.861843	YES

Multiple comparisons: Tukey's HSD using mc.R

#	i	j	n.i	n.j	i-j.diff	lcl	ucl	signif?
# 1	1pg	1vir	25	25	8.04	-3.560486	19.640486	no
# 2	1pg	8pg	25	25	1.44	-10.160486	13.040486	no
# 3	1pg	8vir	25	25	26.08	14.479514	37.680486	YES
# 4	1pg	none	25	25	1.24	-10.360486	12.840486	no
# 5	1vir	8pg	25	25	-6.60	-18.200486	5.000486	no
# 6	1vir	8vir	25	25	18.04	6.439514	29.640486	YES
# 7	1vir	none	25	25	-6.80	-18.400486	4.800486	no
# 8	8pg	8vir	25	25	24.64	13.039514	36.240486	YES
# 9	8pg	none	25	25	-0.20	-11.800486	11.400486	no
# 10	8vir	none	25	25	-24.84	-36.440486	-13.239514	YES

Compare interval lengths for 8v – none

```
Fisher.CI.length <- abs(-33.1326709 + 16.547329)
Bonferroni.CI.length <- abs(-36.818157 + 12.861843)
Tukey.CI.length <- abs(-36.440486 + 13.239514)
CI.lengths <- round(cbind(Fisher=Fisher.CI.length,
                          Bonferroni=Bonferroni.CI.length,
                          Tukey=Tukey.CI.length), 3)
```

```
rownames(CI.lengths) <- "CI.length"
CI.lengths
```

```
##           Fisher Bonferroni Tukey
## CI.length 16.585      23.956 23.201
```

Group statistics.

See “Cannon notes/chap 05/example 5.1-5.6, 5.10-5.11 fruit flies/fruitflies.Rmd”

```
n <- with(data, tapply(Longevity, Treatment, length))
mean <- with(data, round(tapply(Longevity, Treatment, mean), 3))
sd <- with(data, round(tapply(Longevity, Treatment, sd), 3))
idx <- c(5, 1, 3, 2, 4) # idx orders the rows in the table
fruitfly.statistics <- cbind(n, mean, sd)[idx, ]
fruitfly.statistics
```

```
##           n mean    sd
## none      25 63.56 16.452
## 1 pregnant 25 64.80 15.652
## 8 pregnant 25 63.36 14.540
## 1 virgin   25 56.76 14.928
## 8 virgin   25 38.72 12.102
```

Comparison: $\mu_{8v} - \mu_{none} = 0$?

See “Cannon notes/chap 05/example 5.1-5.6, 5.10-5.11 fruit flies/fruitflies.Rmd”

HT:

$H_0 : \mu_{8v} - \mu_{none} = 0$
 $H_a : \mu_{8v} - \mu_{none} \neq 0$

Test statistic:

$$t = \frac{\bar{y}_{8v} - \bar{y}_{none}}{\sqrt{MSE \left(\frac{1}{n_{8v}} + \frac{1}{n_{none}} \right)}}$$

```
y.bar.8v <- fruitfly.statistics[5, 2]
y.bar.none <- fruitfly.statistics[1, 2]
point.estimate <- y.bar.8v - y.bar.none
alpha <- 0.05
df <- 120
mse <- 219.3 # from ANOVA
n.8v <- n.none <- 25
se <- sqrt(mse * (1 / n.8v + 1 / n.none))
t <- point.estimate / se
t
```

```
## [1] -5.930445
```

```
p.value <- 2 * pt(t, df=df)
p.value
```

```
## [1] 2.983554e-08
```

Contrast: $\frac{1}{2}(\mu_{1v} + \mu_{8v}) - \frac{1}{2}(\mu_{1p} + \mu_{8p}) = 0$?

See “Cannon notes/chap 05/example 5.1-5.6, 5.10-5.11 fruit flies/fruitflies.Rmd”

HT:

$H_0 : \frac{1}{2}(\mu_{1v} + \mu_{8v}) - \frac{1}{2}(\mu_{1p} + \mu_{8p}) = 0$

$H_a : \frac{1}{2}(\mu_{1v} + \mu_{8v}) - \frac{1}{2}(\mu_{1p} + \mu_{8p}) \neq 0$

Test statistic:

$$t = \frac{\frac{1}{2}\bar{y}_{1v} + \frac{1}{2}\bar{y}_{8v} - \frac{1}{2}\bar{y}_{1p} - \frac{1}{2}\bar{y}_{8p}}{\sqrt{MSE \cdot \frac{1}{4} \cdot \left(\frac{1}{n_{1v}} + \frac{1}{n_{8v}} + \frac{1}{n_{1p}} + \frac{1}{n_{8p}}\right)}}$$

```
y.bar.1v <- fruitfly.statistics[4, 2]
y.bar.8v <- fruitfly.statistics[5, 2]
y.bar.1p <- fruitfly.statistics[2, 2]
y.bar.8p <- fruitfly.statistics[3, 2]
point.estimate <- (y.bar.1v + y.bar.8v - y.bar.1p - y.bar.8p) / 2
alpha <- 0.05
df <- 120
mse <- 219.3 # from ANOVA
n.1v <- n.8v <- n.1p <- n.8p <- 25
se <- sqrt(mse * (1 / 4) * (1 / n.1v + 1 / n.8v + 1 / n.1p + 1 / n.8p))
t <- point.estimate / se
t
```

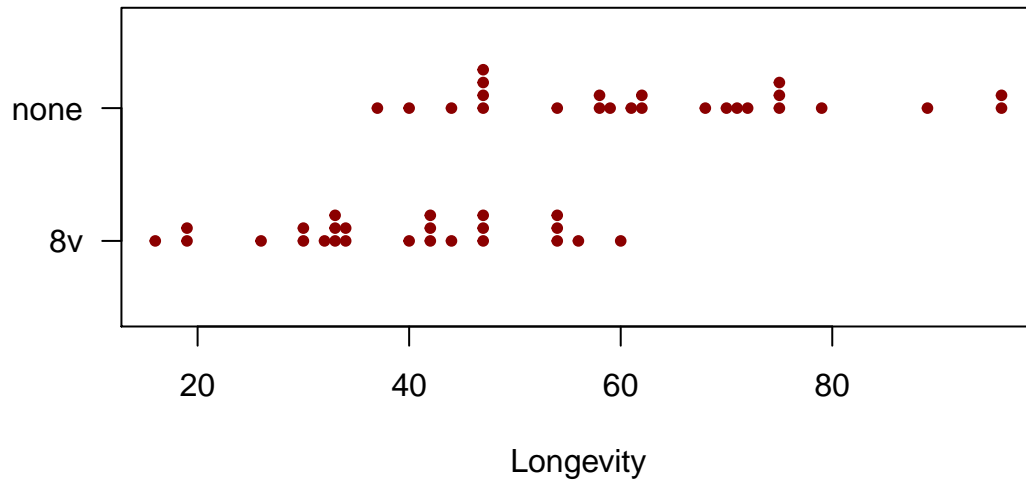
```
## [1] -5.516997
```

```
p.value <- 2 * pt(t, df=df)
p.value
```

```
## [1] 2.017931e-07
```

ANOVA and regression with indicators: pooled two-sample t-test.

```
group.8v <- data[data$Treatment=="8 virgin", "Longevity"]
group.none <- data[data$Treatment=="none", "Longevity"]
stripchart(list(group.8v, group.none),
  pch=20, col="darkred", at=c(1.25, 1.7), method="stack",
  las=1, ylim=1:2, group.names=c("8v", "none"), xlab="Longevity")
```



```
t.test(group.8v, group.none, var.equal=TRUE) # pooled t-test
```

```
##
## Two Sample t-test
##
## data: group.8v and group.none
## t = -6.0811, df = 48, p-value = 1.885e-07
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -33.05298 -16.62702
## sample estimates:
## mean of x mean of y
## 38.72 63.56
```

ANOVA and regression with indicators: ANOVA with two groups.

```
two.groups <- with(data, Treatment=="8 virgin" | Treatment=="none")
fruitfly.aov <- aov(Longevity ~ Treatment, data=data[two.groups, ])
summary(fruitfly.aov)
```

```
##           Df Sum Sq Mean Sq F value    Pr(>F)
## Treatment   1   7713     7713   36.98 1.88e-07
## Residuals  48  10011        209
```

ANOVA and regression with indicators: regression with an indicator.

```
fruitfly.lm <- lm(Longevity ~ Treatment, data=data[two.groups, ])
summary(fruitfly.lm)
```

```
##
## Call:
## lm(formula = Longevity ~ Treatment, data = data[two.groups, ])
##
## Residuals:
##    Min     1Q  Median     3Q    Max
```

```
## -26.56 -8.72 -1.56 8.40 32.44
##
## Coefficients:
##           Estimate Std. Error t value Pr(>|t|)
## (Intercept)   38.720     2.888  13.405 < 2e-16
## Treatmentnone  24.840     4.085   6.081 1.88e-07
##
## Residual standard error: 14.44 on 48 degrees of freedom
## Multiple R-squared:  0.4352, Adjusted R-squared:  0.4234
## F-statistic: 36.98 on 1 and 48 DF,  p-value: 1.885e-07
```

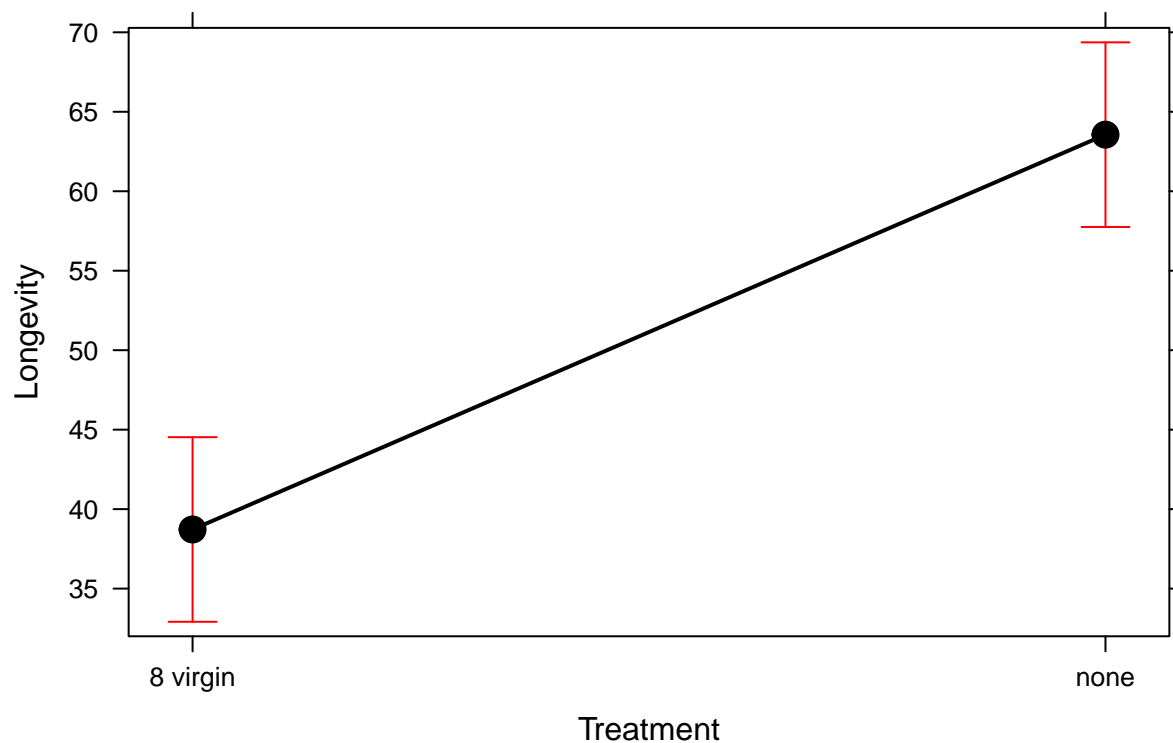
```
confint(fruitfly.lm)
```

```
##           2.5 %   97.5 %
## (Intercept)  32.91255 44.52745
## Treatmentnone 16.62702 33.05298
```

Treatment effects.

```
library(alr4)
plot(allEffects(fruitfly.aov))
```

Treatment effect plot



One-way ANOVA for means as regression: ANOVA


```
fruitfly.aov5 <-aov(Longevity ~ Treatment, data=data) # 5 levels of Treatment
summary(fruitfly.aov5)
```

```
##           Df Sum Sq Mean Sq F value    Pr(>F)
## Treatment    4  11939   2984.8   13.61 3.52e-09
## Residuals  120  26314    219.3
```

One-way ANOVA for means as regression: regression

```
fruitfly.lm5 <-lm(Longevity ~ Treatment, data=data) # 5 levels of Treatment
summary(fruitfly.lm5)
```

```
##
## Call:
## lm(formula = Longevity ~ Treatment, data = data)
##
## Residuals:
##   Min     1Q   Median     3Q      Max
## -35.76  -8.76   0.20  11.20  32.44
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      64.800      2.962  21.880 < 2e-16
## Treatment1 virgin   -8.040      4.188  -1.920  0.0573
## Treatment8 pregnant -1.440      4.188  -0.344  0.7316
## Treatment8 virgin  -26.080      4.188  -6.227 7.27e-09
## Treatmentnone      -1.240      4.188  -0.296  0.7677
##
## Residual standard error: 14.81 on 120 degrees of freedom
## Multiple R-squared:  0.3121, Adjusted R-squared:  0.2892
## F-statistic: 13.61 on 4 and 120 DF,  p-value: 3.516e-09
```

Reorder the factors of Treatment to make “none” the reference.

```
data$Treatment <- factor(data$Treatment,
                          levels=c("none", "1 pregnant", "1 virgin", "8 pregnant", "8 virgin"))
fruitfly.lm5 <-lm(Longevity ~ Treatment, data=data) # 5 levels of Treatment
summary(fruitfly.lm5)
```

```
##
## Call:
## lm(formula = Longevity ~ Treatment, data = data)
##
## Residuals:
##   Min     1Q   Median     3Q      Max
## -35.76  -8.76   0.20  11.20  32.44
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      63.560      2.962  21.461 < 2e-16
## Treatment1 pregnant    1.240      4.188   0.296  0.768
```

```
## Treatment1 virgin    -6.800    4.188   -1.624    0.107
## Treatment8 pregnant  -0.200    4.188   -0.048    0.962
## Treatment8 virgin   -24.840    4.188   -5.931  2.98e-08
##
## Residual standard error: 14.81 on 120 degrees of freedom
## Multiple R-squared:  0.3121, Adjusted R-squared:  0.2892
## F-statistic: 13.61 on 4 and 120 DF,  p-value: 3.516e-09
```

```
aov(fruitfly.lm5)
```

```
## Call:
##   aov(formula = fruitfly.lm5)
##
## Terms:
##           Treatment Residuals
## Sum of Squares  11939.28  26313.52
## Deg. of Freedom      4      120
##
## Residual standard error: 14.80808
## Estimated effects may be unbalanced
```