Homework 6
Matt Massey

Question 1.

```
#read data
irrigation = read.csv("hw6.csv", header = T)

#import packages
library(emmeans)
library(STAT)
library(pwr2)

#get variables from data to make dataframe
method = as.factor(irrigation$METHOD)
block = as.factor(irrigation$BLOCK)
weight = irrigation$WEIGHT
Irrigation = data.frame(method, block, weight)

#model data
Irrigation.lm = lm(weight~ method + block, data = Irrigation)

#means and ci
lsmeans(Irrigation.lm,"method")
##  method lsmean   SE df lower.CL upper.CL
##  1         300 22.9 35      253      346
##  2         290 22.9 35      244      337
##  3         224 22.9 35      177      270
##  4         292 22.9 35      246      338
##  5         291 22.9 35      245      337
##  6         230 22.9 35      183      276
##
## Results are averaged over the levels of: block
## Confidence level used: 0.95


groupmeans = c(300, 290, 224, 292, 291, 230)      #for question 2



#2A ANOVA with blocks
anova(Irrigation.lm)
## Analysis of Variance Table
##
## Response: weight
##           Df Sum Sq Mean Sq F value    Pr(>F)
## method     5  47842    9568  2.2799   0.06783 .
## block      7 457507   65358 15.5733 4.515e-09 ***
## Residuals 35 146888    4197
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#2B Power with blocks
pwr.2way(a = 6, b = 8, alpha = 0.05,
         size.A = 1, size.B = 1, f.A = 0.489, f.B = 0.999)
##
##      Balanced two-way analysis of variance power calculation
##
##                a = 6
##                b = 8
##              n.A = 1
##              n.B = 1
##        sig.level = 0.05
##          power.A = 0.6689196
##          power.B = 0.999118
##            power = 0.6689196
##
## NOTE: power is the minimum power among two factors
```

1A.

This experiment is testing if irrigation method affects orange fruit production (weight of oranges). The null hypothesis states that there is no difference in fruit production from the applied irrigation method. The alternative hypothesis is that there is a difference in fruit production from at least one of the irrigation methods. To achieve a significance level (α) of 0.05, the test statistic and it's corresponding P value of less than 0.05 will cause rejection of the null hypothesis and conclusion that there is at least one irrigation method that significantly alters fruit production. If a P value of 0.05 or greater is obtained, then we will fail to reject the null hypothesis.

The ANOVA model for this experiment gives an F test statistic of 2.2799 and a P value of 0.06783 that is greater than 0.05. Therefore, we fail to reject the null hypothesis and conclude that there is no significant difference in irrigation method on orange fruit production.

1B.

If this experiment were repeated many times under the same conditions, then the hypothesis that all irrigation methods have the same effect on orange fruit production would be rejected 66.9% of the time.

1C.

The key assumption of one-way blocked ANOVA is that there is no interaction between factors and blocks, and all differences between treatment means are assumed constant across blocks. Therefore, if this key assumption is broken or not met, we wouldn't be using the one-way blocked ANOVA to begin with.

Question 2.

```
#2A ANOVA without blocks
Irrigation.lm2 = lm(weight~ method, data = Irrigation)
anova(Irrigation.lm2)
## Analysis of Variance Table
##
## Response: weight
##          Df Sum Sq Mean Sq F value Pr(>F)
## method    5  47842  9568.3  0.6649 0.6521
## Residuals 42 604395 14390.4

oneway.test(weight ~ method, data=Irrigation, var.equal = TRUE)
##
##   One-way analysis of means
##
## data:  weight and method
## F = 0.66491, num df = 5, denom df = 42, p-value = 0.6521

#2B Power without blocks
power.anova.test(groups = 6, n=8, between.var = var(groupmeans), within.var = 14390.3601,
                 sig.level = 0.05, power = NULL)
##
##       Balanced one-way analysis of variance power calculation
##
##          groups = 6
##               n = 8
##     between.var = 1186.567
##      within.var = 14390.36
##       sig.level = 0.05
##           power = 0.2158881
##
## NOTE: n is number in each group

#2D compute sample size
power.anova.test(groups=6, n=NULL, between.var=var(groupmeans), within.var=14390.4,
sig.level=0.05, power=.665)
##
##       Balanced one-way analysis of variance power calculation
##
##          groups = 6
##               n = 24.61021
##     between.var = 1186.567
##      within.var = 14390.4
##       sig.level = 0.05
##           power = 0.665
##
## NOTE: n is number in each group
```

2A.

This experiment is testing if irrigation method affects orange fruit production, and is the same setting as that presented in question 1. The null hypothesis states that there is no difference in fruit production from the applied irrigation method. The alternative hypothesis is that there is a difference in fruit production from at least one of the irrigation methods. To achieve a significance level (α) of 0.05, a P value less than 0.05 will cause rejection of the null hypothesis and conclusion that there is at least one irrigation method that significantly alters fruit production. If a P value of 0.05 or greater is obtained, then we will fail to reject the null hypothesis.

Ignoring the blocks of trees in this example, and only testing for the effects of irrigation methods on fruit production ANOVA for this experiment gives an F test statistic of 0.6649 and a P value of 0.6521 that is greater than 0.05. Therefore, we fail to reject the null hypothesis and conclude that there is no significant difference in irrigation method on orange fruit production.

2B.

If this experiment were repeated many times under the same conditions, then the hypothesis that all irrigation methods have the same effect on orange fruit production would be rejected 21.6% of the time.

2C.

Blocking is a technique to reduce experimental error variance or the effect of nuisance factors, by blocking (grouping) similar experimental units and then considering the effect of factors within each block. In this case, we have removed the blocking of trees so have a much greater experimental error variance.

2D.

To obtain the same power of analysis as including blocks as in question 1 (66.9%), we would need 25 replicates (24.61021 to be exact, but a fractional replicate is not realistic) for each of the 6 treatments.