

# HW7

2024-10-22

STAT 5000 HOMEWORK #7

FALL 2024 DUE FRI, OCTOBER 25TH @ 11:59 PM NAME: SAM OLSON

COLLABORATORS: **The Hatman**

## Q1

For each of the following experiments, determine its

- (i) design: completely randomized, randomized complete block (RCBD), or neither;
- (ii) experimental units, treatments, and blocks (if any blocks are used).
- (iii) Outline an ANOVA table listing sources of variation and degrees of freedom.

### (a)

An agronomist is interested in the effects of plant density on the yield of a certain crop. There are six densities of interest (7, 8, 9, 10, 11, and 12 plants/m<sup>2</sup>) that are interesting. The agronomist has five fields and each field is divided into 6 plots of equal size. She suspects that there is non-homogeneity of soil fertility across fields. She chooses a design such that each density is randomly assigned to one plot in each of the five fields. A separate random assignment is done within each field.

#### i. Design

- Randomized Complete Block Design (**RCBD**): The design uses blocks (five fields with 6 plots each) to account for potential variability in soil fertility across fields, and within each block (plots of a field), the treatments (the six plant densities) are randomly assigned to the plots.

#### ii. Experimental units, treatments, and blocks:

- Experimental units: Plots within fields (6 plots each within one of the 5 fields where the density treatment is applied, totaling 30 plots)
- Treatments: Six plant densities (7, 8, 9, 10, 11, and 12 plants/m<sup>2</sup>)
- Blocks: Five fields (each field serves as a block to account for soil fertility differences)

**iii. ANOVA Table Outline:**

Source of Variation	Degrees of Freedom
Blocks (Fields)	4
Treatments (Densities)	5
Error	$(4)(5) = 20$
Total	29

- Blocks:  $k - 1$ , where  $k = 5$  fields (blocks)
- Treatments:  $t - 1$ , where  $t = 6$  densities (treatments)
- Error:  $(k - 1)(t - 1)$ , where  $k = 5$  fields and  $t = 6$  treatments
- Total: Total number of plots minus one ( $30 - 1 = 29$ )

**(b)**

A router is used to cut notches in a printed circuit board. The vibration of the board as it is cut is considered a major source of variation in the diameter of the notches. Three cutting speeds, 50, 70, and 90 rpm, were used in an experiment to assess the impact of cutting speed on vibration. For each cutting speed, notches were cut in five different circuit boards. Once a notch is cut into a board it cannot be used again, so 15 different boards were used. The assignment of boards to cutting speed was done at random.

**i.**

Design:

- Completely Randomized Design (**CRD**): Since the assignment of boards to the three cutting speeds was done randomly and no blocking is mentioned to account for other sources of variation, this is a completely randomized design, though we have three different treatments being studied.

**ii. Experimental units, treatments, and blocks:**

- Experimental units: Individual circuit boards (15 different boards used in the experiment)
- Treatments: Three cutting speeds (50, 70, and 90 rpm)
- Blocks: No blocks are used in this experiment, as the design is completely randomized.

### iii. ANOVA Table Outline:

Source of Variation	Degrees of Freedom
Treatments (Cutting Speeds)	2
Error	$15 - 3 = 12$
Total	14

- Treatments:  $t - 1$ , where  $t = 3$  cutting speeds (treatments)
- Error: Total experimental units minus number of treatments,  $n - t$ , where  $n = 15$  boards and  $t = 3$  treatments
- Total: Total number of experimental units minus one ( $15 - 1 = 14$ )

### (c)

Industrial psychologists wish to investigate the effect of music in the factory of the productivity of workers. Four distinct music programs and no music make up the five treatments. The experiment is run in 8 plants. Each music program is used for one week. Within each plant music programs are randomly assigned to weeks (1,2,3,4,5) so that all five music programs are used in each plant. Production at each plant is recorded for each of the five weeks.

#### i. Design:

Randomized Complete Block Design (**RCBD**): Since each of the 8 plants serves as a block and all five music programs (treatments) are applied in a randomized order within each plant (across the 5 weeks), this is a randomized complete block design.

#### ii. Experimental units, treatments, and blocks:

- Experimental units: A week of production within a plant (one of the five weeks for a specific plant out of the 8 plants, totaling 40 week-plant combinations)
- Treatments: Five music programs (four distinct music programs and no music)
- Blocks: Eight plants (each plant is a block)

#### iii.

ANOVA Table Outline:

Source of Variation	Degrees of Freedom
Blocks (Plants)	7
Treatments (Music Programs)	4
Error	$(7)(4) = 28$
Total	39

Explanation of Degrees of Freedom: - Blocks:  $k - 1$ , where  $k = 8$  plants (blocks)

- Treatments:  $t - 1$ , where  $t = 5$  music programs (treatments)
- Error:  $(k - 1)(t - 1)$ , where  $k = 8$  plants and  $t = 5$  treatments
- Total: Total number of observations (plants  $\times$  weeks) minus one ( $40 - 1 = 39$ )

## Q2

Standing from a seated position generates stress on the knee joint. Finding ways to minimize this stress is important in individuals who have had knees replaced (Total Knee Arthroplasty, TKA). An experiment is conducted to test whether different feet placement (neutral, back, staggered) affects the amount of torque on the knee (measured in Newton meters, Nm) of 14 older men with TKA. Each of the men will experience all three treatments, in randomly selected order, during the experiment. For all trials, the height of the chair will be the same and participants will all wear comfortable clothing and tennis shoes. The data are located in the file knee.txt (posted on Canvas).

```
dat <- read.table("C:/Users/samue/OneDrive/Desktop/Iowa_State_PS/STAT 5000/PS/PS7/knee.txt") |>
  data.frame(
  )
names(dat) <- c("row", "placement", "torque")
names(dat)
```

```
## [1] "row"      "placement" "torque"
```

Subject	Neutral (N)	Back (B)	Staggered (S)
1	26.3	23.0	21.3
2	22.7	19.7	19.8
3	21.1	20.0	19.5
4	25.9	23.3	22.4
5	25.6	21.1	18.5
6	22.5	19.9	19.6
7	26.1	23.4	22.5
8	21.4	18.8	19.0
9	17.4	17.7	16.6
10	22.8	21.9	20.0
11	23.0	22.2	22.1
12	18.4	17.3	18.6
13	29.6	25.5	21.8
14	27.4	23.2	22.3

(a)

Construct an ANOVA table for the model with fixed placement and subject effects. Include the full table below.

```
placement_means <- tapply(X = dat$torque,
  INDEX = dat$placement,
  FUN = mean)

grand_mean <- mean(dat$torque)

n_j <- table(dat$placement)
SSB <- sum(n_j * (placement_means - grand_mean)^2)
SSB # 81.0876
```

```
## [1] 81.08762
```

```

row_means <- tapply(X = dat$torque,
                    INDEX = dat$row,
                    FUN = mean)

SSR <- sum(table(dat$row) * (row_means - grand_mean)^2)
SSR

```

```
## [1] 231.4924
```

```

SSE <- sum((dat$torque - ave(x = dat$torque,
                           by = dat$placement,
                           FUN = mean)
          - ave(x = dat$torque,
               by = dat$row,
               FUN = mean) + grand_mean)^2)

SSE # 270.6514

```

```
## [1] 39.15905
```

```

SST <- sum((dat$torque - grand_mean)^2)
SST # 351.739

```

```
## [1] 351.739
```

```

df_placement <- length(unique(dat$placement)) - 1
df_row <- length(unique(dat$row)) - 1
df_residual <- nrow(dat) - (df_placement + df_row + 1)
df_total <- nrow(dat) - 1

```

```

MSB_placement <- SSB / df_placement
MSB_row <- SSR / df_row
MSW_residual <- SSE / df_residual

```

```

F_placement <- MSB_placement / MSW_residual
F_row <- MSB_row / MSW_residual

```

```
MSB_placement
```

```
## [1] 40.54381
```

```
F_placement
```

```
## [1] 26.91943
```

```
MSB_row
```

```
## [1] 17.80711
```

F\_row

```
## [1] 11.82319
```

Source	SS	df	MS	F
Between Groups	81.08762	2	40.54381	26.91943
Rows (Block)	231.4924	13	17.80711	11.82319
Residual (Error)	39.15905	26	1.506117	
Total	351.739	41		

Compared to the R output:

```
dat$placement <- as.factor(dat$placement)
dat$row <- as.factor(dat$row)

anova_model <- aov(torque ~ placement + row, data = dat)

summary(anova_model)
```

```
##           Df Sum Sq Mean Sq F value    Pr(>F)
## placement    2  81.09   40.54    26.92 4.63e-07 ***
## row          13 231.49   17.81    11.82 8.53e-08 ***
## Residuals    26  39.16    1.51
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

**(b)**

Are there significant differences in the mean torque between the three feet placements? Conduct an F-test and properly state your conclusion in the context of the study.

There are significant differences in the mean torque between the three feet placements, as we find evidence to reject the null hypothesis at the  $\alpha = 0.01$  level, where the null hypothesis is no difference between the mean feet placements when averaged across the row (subjects). This is evidence in favor of the alternative hypothesis that there is some mean difference in the torque experienced by patients when considering feet placement, when averaging the values of torque across subjects (row).



(c)

Apply Tukey's HSD method with an experiment-wise Type I error level of  $\alpha = 0.05$ . Properly state your conclusions about the pairwise comparison of the mean torque between the three placement positions in the context of the study.

```
tukeyRes <- TukeyHSD(anova_model, "placement", conf.level = 0.95)
print(tukeyRes)
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = torque ~ placement + row, data = dat)
##
## $placement
##           diff          lwr          upr          p adj
## N-B  2.3714286  1.218803  3.524054  0.0000722
## S-B -0.9285714 -2.081197  0.224054  0.1317890
## S-N -3.3000000 -4.452625 -2.147375  0.0000004
```

We have evidence at the  $\alpha = 0.001$  level to reject the null hypothesis in favor of the alternative hypothesis that there is a statistically significant difference in knee torque when using the Neutral foot placement compared to the Back foot placement. The null hypothesis being rejected is that there is no difference in average torque between the neutral and the back placements when averaging torque levels across subjects (rows).

We additionally have evidence at the  $\alpha = 0.001$  level to reject the null hypothesis in favor of the alternative hypothesis that there is a statistically significant difference in knee torque when using the staggered foot placement compared to the neutral foot placement. The null hypothesis being rejected is that there is no difference in average torque between the staggered and the neutral placements when averaging torque levels across subjects (rows).

On the flipside, we do not have evidence to reject the null hypothesis (particularly at the  $\alpha = 0.10$  level) that there is no difference in average torque between the staggered and the back placements when averaging torque levels across subjects (rows).

(d)

Specify and test two orthogonal contrasts that compare (i) the mean torque of the feet neutral group to the mean torque of the other two groups, and (ii) the mean torque of the feet back group to the mean torque of the feet staggered group. How do these results support your conclusions from part (c)?

```
unique(dat$placement)
```

```
## [1] N B S
## Levels: B N S
```

```
dat$placement <- factor(dat$placement, levels = c("N", "B", "S"))
```

```
library(emmeans)
```

```
## Welcome to emmeans.
## Caution: You lose important information if you filter this package's results.
## See '? untidy'
```

```
anova_model_contrasts <- aov(torque ~ placement, data = dat)
```

```
single.means <- emmeans(anova_model_contrasts, "placement")
```

```
# confint(single.means)
test(single.means)
```

```
## placement emmean SE df t.ratio p.value
## N          23.6 0.704 39 33.500 <.0001
## B          21.2 0.704 39 30.131 <.0001
## S          20.3 0.704 39 28.813 <.0001
```

```
contrast(single.means, method=list(
  "n compared to 1/2(b+s)" = c(1, -0.5, -0.5),
  "b compared to s" = c(0, 1, -1)))
```

```
## contrast estimate SE df t.ratio p.value
## n compared to 1/2(b+s) 2.836 0.862 39 3.289 0.0021
## b compared to s 0.929 0.996 39 0.933 0.3568
```

```
unique(dat$placement)
```

```
## [1] N B S
## Levels: N B S
```

```
dat$placement <- factor(dat$placement, levels = c("N", "B", "S"))
```

```
levels(dat$placement)
```

```
## [1] "N" "B" "S"
```

```

contrasts(dat$placement) <- matrix(c(
  1, -0.5, -0.5, # n compared to 1/2(b+s)
  0, 1, -1      # b compared to s
), ncol = 2)

anova_model_contrasts <- aov(torque ~ placement, data = dat)

summary.lm(anova_model_contrasts)

```

```

##
## Call:
## aov(formula = torque ~ placement, data = dat)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -6.1857 -1.4643 -0.3857  2.0143  6.0143
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   21.6952     0.4065  53.372 < 2e-16 ***
## placement1     1.8905     0.5749   3.289  0.00214 **
## placement2     0.4643     0.4978   0.933  0.35677
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.634 on 39 degrees of freedom
## Multiple R-squared:  0.2305, Adjusted R-squared:  0.1911
## F-statistic: 5.842 on 2 and 39 DF,  p-value: 0.006035

```

Please note: The above created contrast estimates using two methods to ensure consistency. Additionally, the “placement” levels are explicitly specified in the above code, which is important to consider as it is not the default order given. I did this to be especially sure and am making this note as others I have discussed this with had different results/conclusions.

My conclusions from the above:

Contrast 1: With a p-value of 0.00214, we have evidence to reject the null hypothesis that there is no difference between the average torque of “neutral” foot position when averaged across subjects (rows) compared to the average of the mean amount of torque from the “back” and “staggered” foot positions. This is evidence in favor of there being a difference between the average torque of “neutral” foot position when averaged across subjects (rows) compared to the average of the mean amount of torque from the “back” and “staggered” foot positions.

Contrast 2: With a p-value of 0.35677, we do not have evidence to reject the null hypothesis that there is no difference between the average torque of “back” foot position when averaged across subjects (rows) compared to the average of the torque from the “staggered” foot positions.

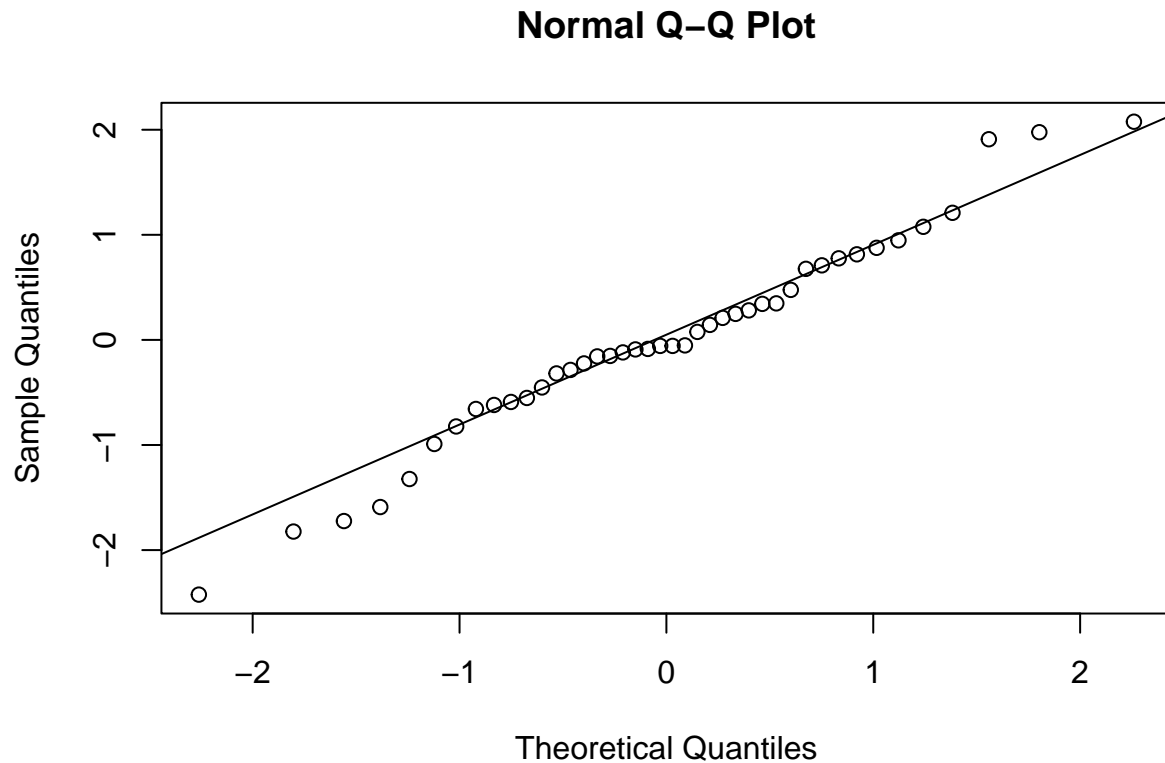
Comparison with Past Analysis: Prior analysis showed that we similarly do not reject the null hypothesis of there being a difference in average torque between the “back” and “staggered” foot positions. Similarly, we found evidence to support the alternative hypotheses that (1) there is a significant difference between the average torque of the “neutral” foot position compared to the “back” foot position, as well (2) there is a significant difference between the average torque of the “neutral” foot position compared to the “staggered” foot position; the first contrast specified is much like a combination of these two tests in one.

(e)

Examine the normal Q-Q plot of the residuals. What does this plot suggest?

```
residuals_anova <- residuals(anova_model)

qqnorm(residuals_anova)
qqline(residuals_anova)
```



The QQ plot of the residuals generally appears to track along its reference line, at least in the center. However, we do observe a general dispersion (lack of fit) compared to the reference line near the tails. Taken together, this is evidence that we likely have normality in our residuals, which means we have evidence to suppose the assumption of normality for ANOVA is not being violated.

(f)

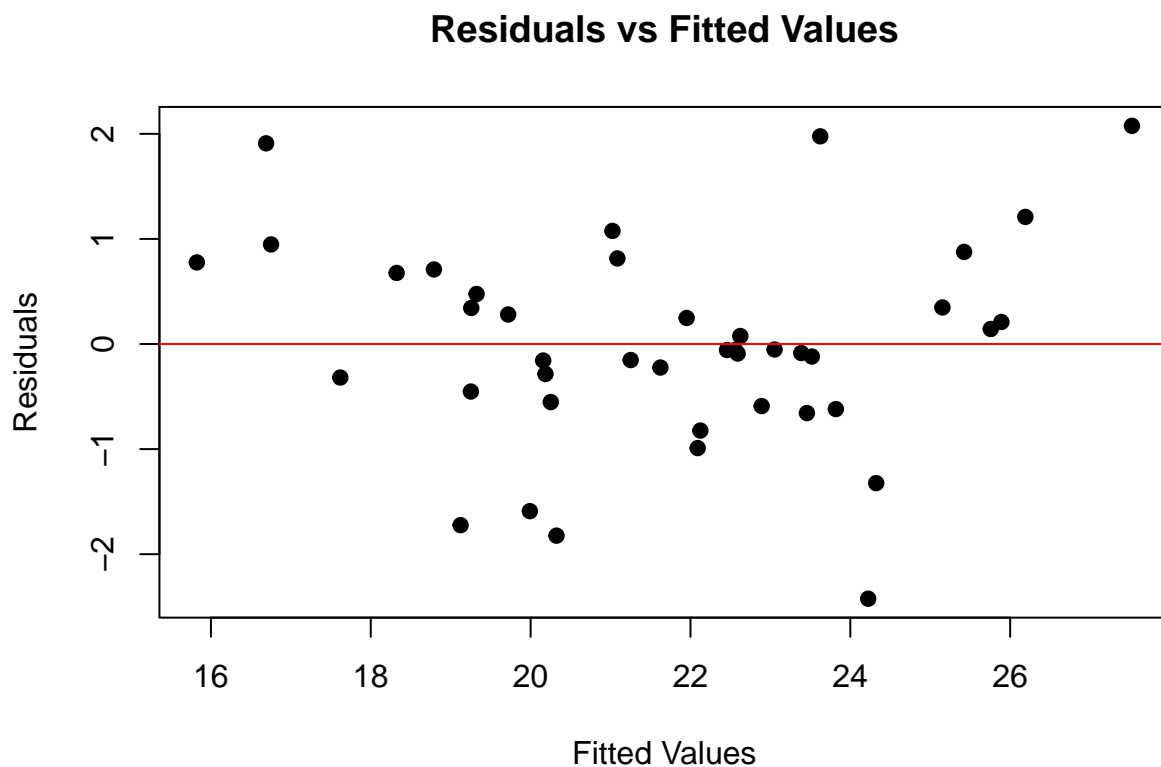
Plot the residuals versus the estimated means (fitted values). What does this plot suggest?

```
residuals_anova <- residuals(anova_model)

fitted_values <- fitted(anova_model)

plot(fitted_values, residuals_anova,
     xlab = "Fitted Values",
     ylab = "Residuals",
     main = "Residuals vs Fitted Values",
     pch = 19) # Adds solid dots for clarity

abline(h = 0, col = "red")
```



Overall, there does not appear to be a general trend in the fitted values compared to the residuals. This indicates to us that we likely do not observe any interaction between the blocks and treatments (between subjects and foot position), and likely are not violating the assumption of additive effects in our model.

### Q3

A group of investigators are interested in the effect of planting density on the yield of millet, a small grain. The investigators are comparing five planting densities, labeled 2, 4, 6, 8, and 10. A field of land was divided into 25 plots (5 rows and 5 columns) to account for known differences in soil quality in the field. In their first study, the investigators randomly assigned density treatments to plots in a 5 x 5 Latin Square. The data are in the millet.txt file (posted on Canvas).

```
millet <- read.table("C:/Users/samue/OneDrive/Desktop/Iowa_State_PS/STAT 5000/PS/PS7/millet.txt",
                    header=TRUE) |>
  data.frame(
  )
```

(a)

Compute the ANOVA table by-hand for the model with fixed row, column, and spacing effects. Include all sources of variation, degrees of freedom, and sums of squares.

```
SS_row <- sum((X = tapply(millet$yield,
                        INDEX = millet$row,
                        FUN = mean) - mean(millet$yield))^2) * nrow(millet)/length(unique(millet$row))

SS_col <- sum((tapply(X = millet$yield,
                    INDEX = millet$col,
                    FUN = mean) - mean(millet$yield))^2) * nrow(millet)/length(unique(millet$col))

SS_spacing <- sum((tapply(X = millet$yield,
                        INDEX = millet$spacing,
                        FUN = mean) - mean(millet$yield))^2) * nrow(millet)/length(unique(millet$spacing))

SS_total <- sum((millet$yield - mean(millet$yield))^2)
SS_residual <- SS_total - SS_row - SS_col - SS_spacing

df_row <- length(unique(millet$row)) - 1
df_col <- length(unique(millet$col)) - 1
df_spacing <- length(unique(millet$spacing)) - 1
df_residual <- nrow(millet) - (df_row + df_col + df_spacing + 1)
df_total <- nrow(millet) - 1

MS_row <- SS_row / df_row
MS_col <- SS_col / df_col
MS_spacing <- SS_spacing / df_spacing
MS_residual <- SS_residual / df_residual

F_row <- MS_row / MS_residual
F_col <- MS_col / MS_residual
F_spacing <- MS_spacing / MS_residual

SS_row
```

```
## [1] 9017.6
```

```
SS_col
```

```
## [1] 4873.2
```

```
SS_spacing
```

```
## [1] 6297.2
```

```
SS_residual
```

```
## [1] 12158
```

```
df_row
```

```
## [1] 4
```

```
df_col
```

```
## [1] 4
```

```
df_spacing
```

```
## [1] 4
```

```
df_residual
```

```
## [1] 12
```

```
MS_row
```

```
## [1] 2254.4
```

```
MS_col
```

```
## [1] 1218.3
```

```
MS_spacing
```

```
## [1] 1574.3
```

```
MS_residual
```

```
## [1] 1013.167
```

F\_row

## [1] 2.225103

F\_col

## [1] 1.202468

F\_spacing

## [1] 1.553841

Source	SS	df	MS	F-value
Rows	9017.6	4	2254.400	2.225103
Columns	4873.2	4	1218.300	1.202468
Spacing	6297.2	4	1574.300	1.553841
Residuals	12158.0	12	1013.167	
Total	32346	24		



(b)

Construct an ANOVA table using R for the model with fixed row, column, and spacing effects to verify your calculations in part (a). Include the full table below.

```
millet$row <- as.factor(millet$row)
millet$col <- as.factor(millet$col)
millet$spacing <- as.factor(millet$spacing)

anova_model <- aov(yield ~ row + col + spacing, data = millet)

summary(anova_model)
```

##	Df	Sum Sq	Mean Sq	F value	Pr(>F)
## row	4	9018	2254	2.225	0.127
## col	4	4873	1218	1.202	0.360
## spacing	4	6297	1574	1.554	0.249
## Residuals	12	12158	1013		

(c)

Are there significant differences in the mean millet yield between the 5 planting densities? Conduct an F-test and properly state your conclusion in the context of the study

```
millet$row <- as.factor(millet$row)
millet$col <- as.factor(millet$col)
millet$spacing <- as.factor(millet$spacing)

anova_model <- aov(yield ~ row + col + spacing, data = millet)

summary(anova_model)
```

```
##           Df Sum Sq Mean Sq F value Pr(>F)
## row         4   9018     2254   2.225  0.127
## col         4   4873     1218   1.202  0.360
## spacing     4   6297     1574   1.554  0.249
## Residuals   12  12158     1013
```

```
# actually just want to use the one from (a)/(b), given here again above
# millet$spacing <- as.factor(millet$spacing)
#
# anova_model_spacing <- aov(yield ~ spacing, data = millet)
#
# summary(anova_model_spacing)
```

Our F-test Null Hypothesis: There is no difference in mean millet yield between the 5 planting densities when controlling for blocking effects of row and column, when averaging across levels of the row and column factors.

Alternative Hypothesis: There is at least one mean difference in millet yield across the 5 planting densities when controlling for blocking effects of row and column, when averaging across levels of the row and column factors.

The above p-value of 0.249 does not provide evidence to reject the null hypothesis stated above. This is consistent with our results from (a) and (b).

In the context of the study, “spacing” is our name for the treatment of the five planting densities. So we do not have evidence to suggest the mean millet yield of the plant densities is significantly different from the other mean millet yields of the other plant densities, when averaging a particular plant densities across our blocking factors of row and column, e.g. the first plant density “2”’s mean millet yield when averaging across row and column is not significantly different than the second plant density “4”’s mean millet yield when averaging across row and column.

(d)

Apply Tukey's HSD method with an experiment-wise Type I error level of  $\alpha = 0.05$ . Properly state your conclusions about the pairwise comparison of the mean millet yield between the five planting densities in the context of the study.

```
millet$row <- as.factor(millet$row)
millet$col <- as.factor(millet$col)
millet$spacing <- as.factor(millet$spacing)

anova_model <- aov(yield ~ row + col + spacing, data = millet)

tukey <- TukeyHSD(anova_model, "spacing", conf.level = 0.95)

tukey
```

```
## Tukey multiple comparisons of means
## 95% family-wise confidence level
##
## Fit: aov(formula = yield ~ row + col + spacing, data = millet)
##
## $spacing
##      diff      lwr      upr      p adj
## 4-2 -16.4 -80.56695 47.76695 0.9210959
## 6-2 -17.4 -81.56695 46.76695 0.9044886
## 8-2 -41.6 -105.76695 22.56695 0.2942426
## 10-2 -40.6 -104.76695 23.56695 0.3149427
## 6-4 -1.0 -65.16695 63.16695 0.9999984
## 8-4 -25.2 -89.36695 38.96695 0.7232595
## 10-4 -24.2 -88.36695 39.96695 0.7504187
## 8-6 -24.2 -88.36695 39.96695 0.7504187
## 10-6 -23.2 -87.36695 40.96695 0.7767384
## 10-8 1.0 -63.16695 65.16695 0.9999984
```

```
# we actually just want the above model
# millet$spacing <- as.factor(millet$spacing)
#
# anovaSpacing <- aov(yield ~ spacing, data = millet)
#
# tukey <- TukeyHSD(anovaSpacing, "spacing", conf.level = 0.95)
#
# tukey
```

All of the above p-values are rather large (all greater than 0.29, and a majority greater than 0.7), such that we do not reject the null hypothesis that the difference in the average millet yield of any pairwise combinations of spacing levels are significantly different from one another when averaging across rows and columns.

Consistent with the above, we are 95% confident that the true difference between any pairwise combination of the spacing levels contains zero (is between a positive and a negative number), which indicates that we are not certain, and do not have evidence to reject, that the null hypothesis that the average millet yield of any pairwise combinations of spacing levels are significantly different from one another when averaging across rows and columns.

(e)

If the study is repeated next year, would you recommend that it still be a Latin Squares design? Or, should the investigators just use row blocks, just use column blocks, or just use a CRD? Justify your answer with appropriate calculations and explanations.

```
millet$row <- as.factor(millet$row)
millet$col <- as.factor(millet$col)
millet$spacing <- as.factor(millet$spacing)

anova_model <- aov(yield ~ row + spacing + col, data = millet)

summary(anova_model)
```

```
##           Df Sum Sq Mean Sq F value Pr(>F)
## row         4   9018     2254   2.225  0.127
## spacing     4   6297     1574   1.554  0.249
## col         4   4873     1218   1.202  0.360
## Residuals   12  12158     1013
```

```
anova_model <- aov(yield ~ row, data = millet)
summary(anova_model)
```

```
##           Df Sum Sq Mean Sq F value Pr(>F)
## row         4   9018     2254   1.933  0.144
## Residuals   20  23328     1166
```

```
anova_model <- aov(yield ~ spacing, data = millet)
summary(anova_model)
```

```
##           Df Sum Sq Mean Sq F value Pr(>F)
## spacing     4   6297     1574   1.209  0.338
## Residuals   20  26049     1302
```

```
anova_model <- aov(yield ~ col, data = millet)
summary(anova_model)
```

```
##           Df Sum Sq Mean Sq F value Pr(>F)
## col         4   4873     1218   0.887  0.49
## Residuals   20  27473     1374
```

```
anova_model <- aov(yield ~ row + spacing, data = millet)
summary(anova_model)
```

```
##           Df Sum Sq Mean Sq F value Pr(>F)
## row         4   9018     2254   2.118  0.126
## spacing     4   6297     1574   1.479  0.255
## Residuals   16  17031     1064
```

```
anova_model <- aov(yield ~ row + col, data = millet)
summary(anova_model)
```

```
##           Df Sum Sq Mean Sq F value Pr(>F)
## row        4   9018    2254    1.954  0.15
## col        4   4873    1218    1.056  0.41
## Residuals  16  18455    1154
```

```
anova_model <- aov(yield ~ spacing + col, data = millet)
summary(anova_model)
```

```
##           Df Sum Sq Mean Sq F value Pr(>F)
## spacing    4   6297    1574    1.190  0.353
## col        4   4873    1218    0.921  0.476
## Residuals  16  21176    1324
```

```
anovaCol <- aov(yield ~ col, data = millet)

tukeyCol <- TukeyHSD(anovaCol, "col", conf.level = 0.95)

tukeyCol
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = yield ~ col, data = millet)
##
## $col
##      diff      lwr      upr      p adj
## 2-1 28.2 -41.94269  98.34269 0.7496627
## 3-1 31.6 -38.54269 101.74269 0.6657835
## 4-1 40.2 -29.94269 110.34269 0.4476250
## 5-1 34.0 -36.14269 104.14269 0.6041425
## 3-2  3.4 -66.74269  73.54269 0.9998910
## 4-2 12.0 -58.14269  82.14269 0.9851795
## 5-2  5.8 -64.34269  75.94269 0.9990972
## 4-3  8.6 -61.54269  78.74269 0.9958094
## 5-3  2.4 -67.74269  72.54269 0.9999728
## 5-4 -6.2 -76.34269  63.94269 0.9988269
```

```
millet$row <- as.factor(millet$row)

anovaRow <- aov(yield ~ row, data = millet)

tukeyRow <- TukeyHSD(anovaRow, "row", conf.level = 0.95)

tukeyRow
```

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

```
## Fit: aov(formula = yield ~ row, data = millet)
##
## $row
##      diff      lwr      upr      p adj
## 2-1  11.0 -53.635845  75.63584 0.9854660
## 3-1   -8.0 -72.635845  56.63584 0.9956542
## 4-1 -16.4 -81.035845  48.23584 0.9392655
## 5-1  38.4 -26.235845 103.03584 0.4128253
## 3-2 -19.0 -83.635845  45.63584 0.9010574
## 4-2 -27.4 -92.035845  37.23584 0.7125543
## 5-2  27.4 -37.235845  92.03584 0.7125543
## 4-3   -8.4 -73.035845  56.23584 0.9947583
## 5-3  46.4 -18.235845 111.03584 0.2394205
## 5-4  54.8  -9.835845 119.43584 0.1217960
```

### **Making a Choice:**

I recommend continuing with the Latin Squares design, as it provides the greatest efficiency (precision of estimates of treatment and block effects), despite making implementation somewhat more complicated for researchers.

In our study, we want to maximize the amount of variability in our Model Sum of Squares, which includes both treatment and blocking effects; this is also to say that we want to minimize the Error (Residual) Sum of Squares.

Despite not being statistically significant from zero for either blocking effect (for a given block, row or column, we do not have evidence to reject the null hypothesis that the mean millet yields are different from any other level of the blocking factor when averaged across treatments and the other respective blocking factor), we still improve the amount of variability (observe greater sum of squares for the Model). So by contrast, models like RCBD (with one blocking factor, i.e. only row or only column) and RCD would (and do in the above) have greater Sum of Squares Error, meaning they are less efficient/less precise than the estimates produced by the Latin Squares model, which uses two blocking factors. It is important to note that this result holds despite not producing any estimates of treatment or block effects that are significantly different from zero (do not reject the null hypothesis as stated previously).

Understandably though, the Latin Squares Design approach is also a bit more difficult for researchers to implement compared to the alternatives. Despite this, I would argue that the improved efficiency would be worth continuing to utilize this design; while certainly a bit of a hand-wave, the researchers would regardless be implementing their treatments via plots on a row-column grid, so it is an inherent part of the research design and worthwhile to incorporate into their model, i.e. to “control” for when analyzing treatment effects.

### **Why not use an RCBD?:**

Well, semantically, Latin Squares is specific instance of an RCBD. That notwithstanding, the question is really why use Latin Squares with the two blocking factors, row and column, rather than just using one blocking factor (just row, or just column)?

Aside from the fact that researchers are implementing treatments based on a combination of row and column (making the choice of “controlling for just row effects” or “controlling for just column effects” rather nonsensical), there are also advantages to be had in terms of precision. Despite not having evidence to reject the respective null hypotheses of either the row or column blocking factors, we nonetheless see that the Latin Squares ANOVA table has a smaller Residual Sum of Squares. This is a key distinction to make, inasmuch as it is evidence that our Latin Squares model provides more precision/efficiency than a single block variable RCBD model.

### **Why not use CRD?:**

At the risk of redundancy, the above assessment can be extended to comparing RCBD/Latin Squares to CRD, including the implications to efficiency gains from the Latin Squares design. From past work and lectures,

we know generally that RCBD has greater efficiency than CRD, and since Latin Squares is a particular type of RCBD, transitively it holds that we would prefer Latin Squares to CRD as it is more efficient.



Figure 1: CocoMelon