FA9 BORROMEO MAYO MERCADO RMD FILE

Borromeo Mayo Mercado

2025-11-22

Introduction

A study was conducted to determine whether crop yield is influenced by (1) fertilizer blend (X, Y, Z) and (2) crop type (wheat, corn, soy, rice). A two-way ANOVA was used to test for:

- A main effect of fertilizer blend,

- A main effect of crop type,

- An interaction between fertilizer blend $\times$ crop type

Assumptions

When you choose to analyse your data using a two-way ANOVA, part of the process involves checking to make sure that the data you want to analyse can actually be analysed using a two-way ANOVA. You need to do this because it is only appropriate to use a two-way ANOVA if your data "passes" six assumptions that are required for a two-way ANOVA to give you a valid result.

Below are the assumptions:

Assumption #1: Your **dependent variable** should be measured at the **continuous** level (i.e., they are **interval** or **ratio** variables).

Assumption #2: Your **two independent variables** should each consist of **two or more categorical**, **independent groups**.

Assumption #3: You should have **independence of observations**, which means that there is no relationship between the observations in each group or between the groups themselves.

Assumption #4: There should be **no significant outliers**.

Assumption #5: Your **dependent variable** should be **approximately normally distributed for each combination of the groups of the two independent variables**.

Assumption #6: There needs to be **homogeneity of variances for each combination of the groups of the two independent variables**.

Null hypothesis:

There is no significant interaction effect on yield between fertilizer and crop.

Problem:

A new fertilizer has been developed to increase the yield on crops, and the makers of the fertilizer want to better understand which of the three formulations (blends) of this fertilizer are most effective for wheat, corn, soybeans and rice (crops). They test each of the three blends on 5 samples of each of the four types of crops. The crop yields for the 12 combinations are as shown in the table below.

```
df %>%
kable(caption = "Crop Yield Data (Long Format)", digits = 0) %>%
kable_styling(bootstrap_options = c("striped", "hover", "condensed"), full_width = FALSE)
```

Table 1: Crop Yield Data (Long Format)

Fertilizer	Crop	Yield
Blend X	Wheat	123
Blend X	Corn	128
Blend X	Soy	166
Blend X	Rice	151
Blend X	Wheat	156
Blend X	Corn	150
Blend X	Soy	178
Blend X	Rice	125
Blend X	Wheat	112
Blend X	Corn	174
Blend X	Soy	187
Blend X	Rice	117
Blend X	Wheat	100
Blend X	Corn	116
Blend X	Soy	153
Blend X	Rice	155
Blend X	Wheat	168
Blend X	Corn	109
Blend X	Soy	195
Blend X	Rice	158
Blend Y	Wheat	135
Blend Y	Corn	175
Blend Y	Soy	140
Blend Y	Rice	167
Blend Y	Wheat	130
Blend Y	Corn	132
Blend Y	Soy	145
Blend Y	Rice	183
Blend Y	Wheat	176
Blend Y	Corn	120
Blend Y	Soy	159
Blend Y	Rice	142
Blend Y	Wheat	120
Blend Y	Corn	187
Blend Y	Soy	131

|		
Blend Y	Rice	167
Blend Y	Wheat	155
Blend Y	Corn	184
Blend Y	Soy	126
Blend Y	Rice	168
Blend Z	Wheat	156
Blend Z	Corn	186
Blend Z	Soy	185
Blend Z	Rice	175
Blend Z	Wheat	180
Blend Z	Corn	138
Blend Z	Soy	206
Blend Z	Rice	173
Blend Z	Wheat	147
Blend Z	Corn	178
Blend Z	Soy	188
Blend Z	Rice	154
Blend Z	Wheat	146
Blend Z	Corn	176
Blend Z	Soy	165
Blend Z	Rice	191
Blend Z	Wheat	193
Blend Z	Corn	190
Blend Z	Soy	188
Blend Z	Rice	169

Assumptions Check

Assumption #1: The dependent variable, crop yield, is measured at the continuous level.

Assumption #2: The two independent variables, fertilizer type (Blend X, Blend Y, Blend Z) and crop type (Wheat, Corn, Soy, Rice), each consist of two or more categorical, independent groups.

Assumption #3: The independence of observation is observed.

Assumption #4: There are no significant outliers in each of the 12 cells of the design.

Raincloud Plot of Crop Yield by Fertilizer
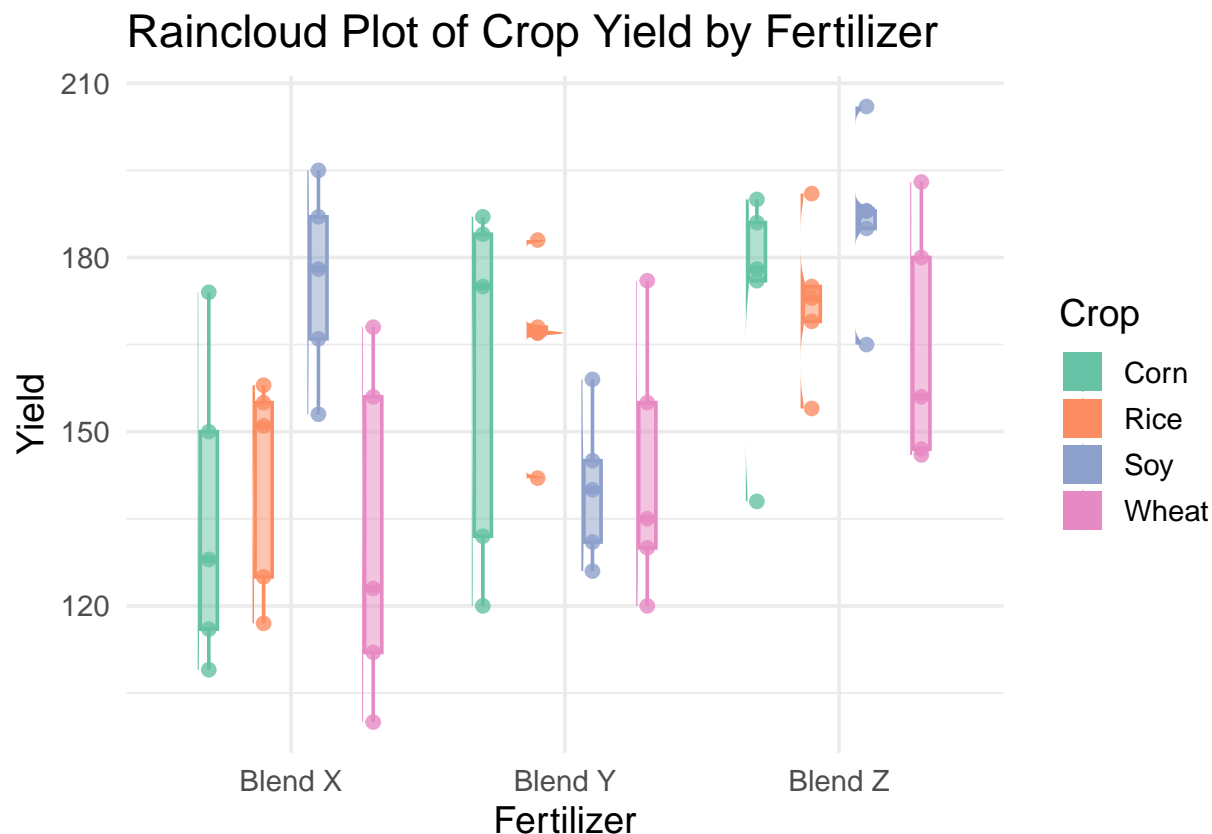
Assumption #5: The dependent variable, crop yield, is approximately normally distributed for each combination of the groups of fertilizer type and crop type, as assessed by Shapiro–Wilk test of normality (p > .05).

Table 2: Descriptive Statistics of Yield by Fertilizer

Statistic	Blend X	Blend Y	Blend Z
Valid	20.000	20.000	20.000
Mean	146.050	152.100	174.200
SD	27.971	22.606	18.185
Skewness	-0.010	0.090	-0.410
SE_Skew	0.548	0.548	0.548
Kurtosis	-1.328	-1.563	-0.907
SE_Kurt	1.095	1.095	1.095
Shapiro_Wilk	0.952	0.924	0.952

Table 3: Descriptive Statistics of Yield by Crop

Statistic	Corn	Rice	Soy	Wheat
Valid	15.000	15.000	15.000	15.000
Mean	156.200	159.667	167.467	146.467
SD	29.513	20.123	24.617	26.568
Skewness	-0.296	-0.607	-0.206	-0.008

SE__Skew	0.632	0.632	0.632	0.632
Kurtosis	-1.719	-0.468	-1.369	-1.139
SE__Kurt	1.265	1.265	1.265	1.265
Shapiro__Wilk	0.867	0.946	0.952	0.983

Assumption #6: The variances for each combination of the groups of fertilizer type and crop type are homogeneous, as assessed by Levene's test of equality of variances, p = 0.755.

Table 4: Test for Equality of Variances (Levene's Test)

df1	df2	F	p
11	48	0.675	0.755

Computation

```r
library(car)
library(effectsize)
```

```
## Warning: package 'effectsize' was built under R version 4.4.3
```

```
##
## Attaching package: 'effectsize'
```

```
## The following object is masked from 'package:psych':
##
##     phi
```

```
## The following objects are masked from 'package:rstatix':
##
##     cohens_d, eta_squared
```

```r
library(kableExtra)
library(tidyverse)

# Type III ANOVA
anova_res <- Anova(lm(Yield ~ Fertilizer * Crop, data = df), type = 3)

# Convert ANOVA table to data frame
anova_df <- as.data.frame(anova_res) %>%
  rownames_to_column(var = "Cases")

# Rename columns
anova_table <- anova_df %>%
  rename(
    `Sum of Squares` = `Sum Sq`,
    df = `Df`,
    F = `F value`,
    p = `Pr(>F)`
```

```r
  ) %>%
  mutate(
    `Mean Square` = `Sum of Squares` / df
  )

# Partial Eta Squared
SS_error <- anova_table$`Sum of Squares`[anova_table$Cases == "Residuals"]
anova_table$`Partial Eta Squared` <- anova_table$`Sum of Squares` /
  (anova_table$`Sum of Squares` + SS_error)

# Clean Cases names
anova_table$Cases <- gsub("\\(|\\)", "", anova_table$Cases)

# Arrange columns
anova_table <- anova_table %>%
  select(Cases, `Sum of Squares`, df, `Mean Square`, F, p, `Partial Eta Squared`)

# Print LaTeX table
kable(
  anova_table,
  caption = "Two-way ANOVA Table (JASP-style, Type III SS)",
  digits = 3,
  format = "latex",
  booktabs = TRUE
) %>%
  kable_styling(latex_options = c("hold_position"), full_width = FALSE)
```

Table 5: Two-way ANOVA Table (JASP-style, Type III SS)

Cases	Sum of Squares	df	Mean Square	F	p	Partial Eta Squared
Intercept	91665.80	1	91665.800	207.346	0.000	0.812
Fertilizer	3734.80	2	1867.400	4.224	0.020	0.150
Crop	6125.35	3	2041.783	4.618	0.006	0.224
Fertilizer:Crop	6225.90	6	1037.650	2.347	0.046	0.227
Residuals	21220.40	48	442.092	NA	NA	0.500

Remark:

Results of Simple Main Effects

```r
library(rstatix)
library(dplyr)
library(kableExtra)

overall_aov <- aov(Yield ~ Fertilizer * Crop, data = df_long)
MS_Error <- summary(overall_aov)[[1]][3, "Mean Sq"]

simple_effects_fert <- df_long %>%
```

```
  group_by(Crop) %>%
  anova_test(Yield ~ Fertilizer) %>%
  get_anova_table() %>%
  as_tibble() %>%
  select(Crop, Effect, DFn, DFd, F, p) %>%
  rename(`Level of Crop` = Crop,
         `Source` = Effect,
         df = DFn) %>%
  mutate(
    `Mean Square` = F * MS_Error,
    `Sum of Squares` = `Mean Square` * df
  ) %>%
  select(`Level of Crop`, `Sum of Squares`, df, `Mean Square`, F, p)

simple_effects_fert %>%
  kable(
    caption = "Simple Main Effects of Fertilizer within Each Crop",
    digits = 3
  ) %>%
  kable_styling(bootstrap_options = c("striped", "hover", "condensed"), full_width = FALSE)
```

Table 6: Simple Main Effects of Fertilizer within Each Crop

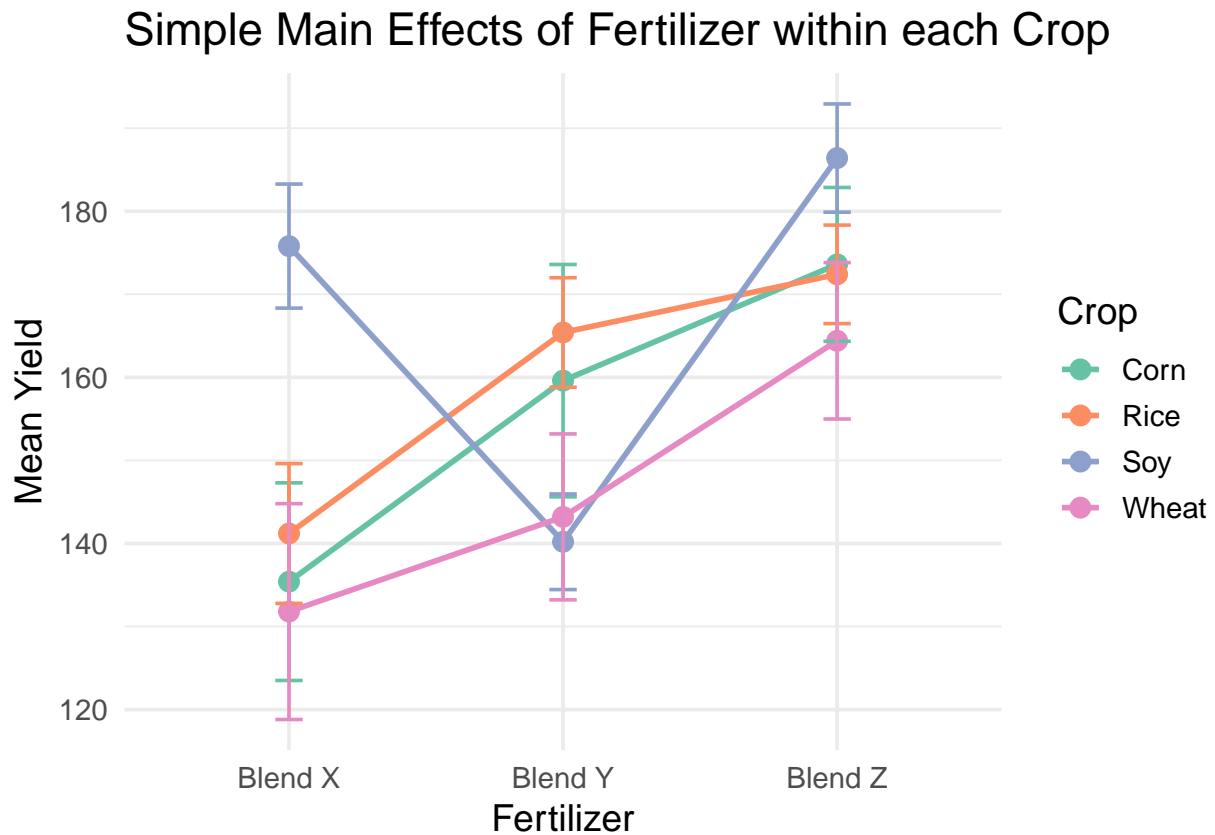| Level of Crop	Sum of Squares	df	Mean Square	F	p
Corn	5497.470	2	2748.735	2.649	0.111
Rice	11165.114	2	5582.557	5.380	0.021
Soy	27763.363	2	13881.682	13.378	0.001
Wheat	4769.039	2	2384.520	2.298	0.143

```
sme_means <- df %>%
group_by(Fertilizer, Crop) %>%
summarise(
Mean_Yield = mean(Yield),
SD = sd(Yield),
N = n(),
SE = SD/sqrt(N),
.groups = "drop"
)

ggplot(sme_means, aes(x = Fertilizer, y = Mean_Yield, group = Crop, color = Crop)) +
geom_line(size = 1) +
geom_point(size = 3) +
geom_errorbar(aes(ymin = Mean_Yield - SE, ymax = Mean_Yield + SE), width = 0.1) +
theme_minimal(base_size = 14) +
labs(
title = "Simple Main Effects of Fertilizer within each Crop",
x = "Fertilizer",
y = "Mean Yield"
) +
scale_color_brewer(palette = "Set2") +
theme(legend.position = "right")
```

```
## Warning: Using `size` aesthetic for lines was deprecated in ggplot2 3.4.0.
## i Please use `linewidth` instead.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.
```


Simple Main Effects of Fertilizer within each Crop

Remark:

```
simple_effects_crop <- df_long %>%
  group_by(Fertilizer) %>%
  anova_test(Yield ~ Crop) %>%
  get_anova_table() %>%
  as_tibble() %>%
  select(Fertilizer, Effect, DFn, DFd, F, p) %>%
  rename(`Level of Fertilizer` = Fertilizer,
         `Source` = Effect,
         df = DFn) %>%
  mutate(
    `Mean Square` = F * MS_Error,
    `Sum of Squares` = `Mean Square` * df
  ) %>%
  select(`Level of Fertilizer`, `Sum of Squares`, df, `Mean Square`, F, p)
```

```
simple_effects_crop %>%
  kable(
    caption = "Simple Main Effects of Crop within Each Fertilizer",
    digits = 3
  ) %>%
  kable_styling(bootstrap_options = c("striped", "hover", "condensed"), full_width = FALSE)
```

Table 7: Simple Main Effects of Crop within Each Fertilizer

| Level of Fertilizer	Sum of Squares	df	Mean Square	F	p
Blend X	11636.207	3	3878.736	3.738	0.033
Blend Y	5064.770	3	1688.257	1.627	0.223
Blend Z	4093.529	3	1364.510	1.315	0.304

```
sme_means2 <- df %>%
group_by(Crop, Fertilizer) %>%
summarise(
Mean_Yield = mean(Yield),
SD = sd(Yield),
N = n(),
SE = SD/sqrt(N),
.groups = "drop"
)

ggplot(sme_means2, aes(x = Crop, y = Mean_Yield, group = Fertilizer, color = Fertilizer)) +
geom_line(size = 1) +
geom_point(size = 3) +
geom_errorbar(aes(ymin = Mean_Yield - SE, ymax = Mean_Yield + SE), width = 0.1) +
theme_minimal(base_size = 14) +
labs(
title = "Simple Main Effects of Crop within each Fertilizer",
x = "Crop",
y = "Mean Yield"
) +
scale_color_brewer(palette = "Set2") +
theme(legend.position = "right")
```

Simple Main Effects of Crop within each Fertilizer
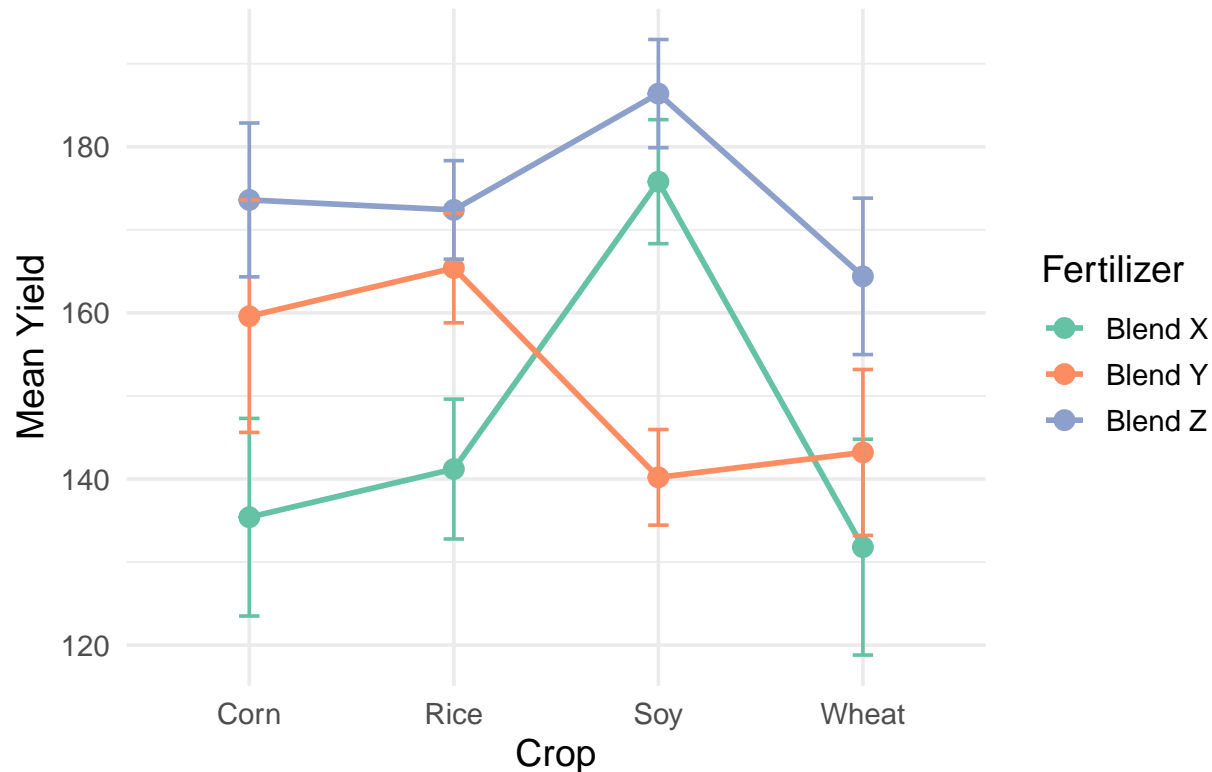
Remark:

Partial Eta Squared

```r
eta_sq_res <- eta_squared(
  lm(Yield ~ Fertilizer * Crop, data = df),
  partial = TRUE
)

eta_sq_res %>%
  rename(Effect = Parameter,
         Partial_Eta_Sq = Eta2_partial) %>%
  kable(caption = "Partial Eta Squared for Two-way ANOVA",
        digits = 3) %>%
  kable_styling(bootstrap_options = c("striped","hover","condensed"),
                full_width = FALSE)
```

Table 8: Partial Eta Squared for Two-way ANOVA

Effect	Partial_Eta_Sq	CI	CI_low	CI_high
Fertilizer	0.293	0.95	0.113	1
Crop	0.139	0.95	0.000	1

Fertilizer:Crop	0.227	0.95	0.003	1

Post Hoc Comparisons

```r
library(emmeans)
```

```
## Warning: package 'emmeans' was built under R version 4.4.3
```

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

```r
library(kableExtra)
library(dplyr)

# Fit two-way ANOVA model
model <- lm(Yield ~ Fertilizer * Crop, data = df_long)

# --- Estimated Marginal Means ---
emm_fert <- emmeans(model, ~ Fertilizer)
```

```
## NOTE: Results may be misleading due to involvement in interactions
```

```r
emm_crop <- emmeans(model, ~ Crop)
```

```
## NOTE: Results may be misleading due to involvement in interactions
```

```r
emm_interaction <- emmeans(model, ~ Fertilizer * Crop)

# --- Pairwise Comparisons ---

# 1. Fertilizer main effect
pairwise_fert <- contrast(emm_fert, method = "pairwise", adjust = "tukey") %>% summary(infer = TRUE)
pairwise_fert_df <- as.data.frame(pairwise_fert)

# Print Fertilizer post hoc table
pairwise_fert_df %>%
  kable(
    caption = "Post Hoc Pairwise Comparisons of Fertilizer (Tukey)",
    digits = 3,
    format = "latex",
    booktabs = TRUE
  ) %>%
  kable_styling(latex_options = c("hold_position"), full_width = FALSE)
```

Table 9: Post Hoc Pairwise Comparisons of Fertilizer (Tukey)

contrast	estimate	SE	df	lower.CL	upper.CL	t.ratio	p.value
Blend X - Blend Y	-6.05	6.649	48	-22.131	10.031	-0.910	0.637
Blend X - Blend Z	-28.15	6.649	48	-44.231	-12.069	-4.234	0.000
Blend Y - Blend Z	-22.10	6.649	48	-38.181	-6.019	-3.324	0.005

```
# 2. Crop main effect
pairwise_crop <- contrast(emm_crop, method = "pairwise", adjust = "tukey") %>% summary(infer = TRUE)
pairwise_crop_df <- as.data.frame(pairwise_crop)

# Print Crop post hoc table
pairwise_crop_df %>%
  kable(
    caption = "Post Hoc Pairwise Comparisons of Crop (Tukey)",
    digits = 3,
    format = "latex",
    booktabs = TRUE
  ) %>%
  kable_styling(latex_options = c("hold_position"), full_width = FALSE)
```

Table 10: Post Hoc Pairwise Comparisons of Crop (Tukey)

contrast	estimate	SE	df	lower.CL	upper.CL	t.ratio	p.value
Corn - Rice	-3.467	7.678	48	-23.900	16.966	-0.452	0.969
Corn - Soy	-11.267	7.678	48	-31.700	9.166	-1.467	0.465
Corn - Wheat	9.733	7.678	48	-10.700	30.166	1.268	0.588
Rice - Soy	-7.800	7.678	48	-28.233	12.633	-1.016	0.741
Rice - Wheat	13.200	7.678	48	-7.233	33.633	1.719	0.325
Soy - Wheat	21.000	7.678	48	0.567	41.433	2.735	0.042

```
# 3. Interaction: Fertilizer pairwise within each Crop
pairwise_interaction <- contrast(emm_interaction, method = "pairwise", by = "Crop", adjust = "tukey") %
pairwise_interaction_df <- as.data.frame(pairwise_interaction)

# Print Interaction post hoc table
pairwise_interaction_df %>%
  kable(
    caption = "Post Hoc Pairwise Comparisons of Fertilizer within Each Crop (Tukey)",
    digits = 3,
    format = "latex",
    booktabs = TRUE
  ) %>%
  kable_styling(latex_options = c("hold_position"), full_width = FALSE)
```

Results:

Table 11: Post Hoc Pairwise Comparisons of Fertilizer within Each Crop (Tukey)

contrast	Crop	estimate	SE	df	lower.CL	upper.CL	t.ratio	p.value
Blend X - Blend Y	Corn	-24.2	13.298	48	-56.361	7.961	-1.820	0.174
Blend X - Blend Z	Corn	-38.2	13.298	48	-70.361	-6.039	-2.873	0.016
Blend Y - Blend Z	Corn	-14.0	13.298	48	-46.161	18.161	-1.053	0.548
Blend X - Blend Y	Rice	-24.2	13.298	48	-56.361	7.961	-1.820	0.174
Blend X - Blend Z	Rice	-31.2	13.298	48	-63.361	0.961	-2.346	0.059
Blend Y - Blend Z	Rice	-7.0	13.298	48	-39.161	25.161	-0.526	0.859
Blend X - Blend Y	Soy	35.6	13.298	48	3.439	67.761	2.677	0.027
Blend X - Blend Z	Soy	-10.6	13.298	48	-42.761	21.561	-0.797	0.707
Blend Y - Blend Z	Soy	-46.2	13.298	48	-78.361	-14.039	-3.474	0.003
Blend X - Blend Y	Wheat	-11.4	13.298	48	-43.561	20.761	-0.857	0.670
Blend X - Blend Z	Wheat	-32.6	13.298	48	-64.761	-0.439	-2.451	0.046
Blend Y - Blend Z	Wheat	-21.2	13.298	48	-53.361	10.961	-1.594	0.258