Experimental Design Project

STA2005S - 2025

Maryam Abrahams (ABRMAR043)
Bheka Mabika (MBKBHE002)

Table of contents

Plagiarism Declaration:

1. I know that plagiarism is a serious form of academic dishonesty.
2. I have read the document about avoiding plagiarism, am familiar with its contents and have avoided all forms of plagiarism mentioned there.
3. Where I have used the words of others, I have indicated this by the use of quotation marks.
4. I have referenced all quotations and other ideas borrowed from others.
5. I have not and shall not allow others to plagiarise my work.

Signature:

Plagiarism Declaration:

1. I know that plagiarism is a serious form of academic dishonesty.
2. I have read the document about avoiding plagiarism, am familiar with its contents and have avoided all forms of plagiarism mentioned there.
3. Where I have used the words of others, I have indicated this by the use of quotation marks.
4. I have referenced all quotations and other ideas borrowed from others.
5. I have not and shall not allow others to plagiarise my work.

Signature:

Introduction:

The two chilli farmers, Ms Hopeful and Mr Growing, compete head to head for the hearts and minds of the South African chilli market. But despite their names Ms Hopeful seems to always get higher chilli yields than Mr Growing. In Mr Growing's eyes this cannot continue.

Aim:

This experiment aims to systematically evaluate the effect of light levels, heat levels and chilli variety on crop quality and yield. With our results we will be able to give evidence-based recommendations on optimal growing conditions.

Particularly we seek to answer:

1. Which light and heat settings produce the highest quality chillies?
2. Does the above depend on variety?
3. Does quality depend on variety?
4. How big are the differences?

Priori Hypothesis:

Prior to collecting the data, we hypothesize:

- H_1: Mean quality will increase with an increase in light levels
- H_2: Mean Quality will increase with intermediate heat levels (2-3), as extreme heat may damage plant cells while insufficient heat limits growth.
- H_3: There may (or may not) be a Heat × Light interaction.
- H_4: Both varieties will show similar response patterns to light and heat treatments. But mean quality will differ by variety.
- H_5: Plot side (North/South) will have negligible effect on quality after accounting for other factors.

Experimental Design and Randomization:

Simulating data:

We created the design matrix in accordance to a **Nested Latin Square design**, testing all 16 heat × light combinations in each block where each block representd a greenhouse-season combination and is created from a 2$×$2 latin square design with greenhouse (A, B) as columns and season (1, 2) as rows.

treatment factors:

- Light: 4 levels (1, 2, 3, 4)
- Heat: 4 levels (1, 2, 3, 4)

$\implies$ 16 unique treatments

blocking factors:

- Season: (1, 2)
- Greenhouse: (A, B)
- Side of greenhouse: (N = north-facing, S = south-facing)
- Chilli variety: (R = red-hot, F = furious)

response variables: quality score, which combines taste, yield, and look

Side-Variety combination assignments:

	H1	H2	H3	H4
L1	RN	FN	RS	FS
L2	FN	RN	FS	RS
L3	RS	FS	RN	FN
L4	FS	RS	FN	RN

->

Example Greenhouse setup:

	R		F	
N	H1	L1	H1	L2
	H2	L2	H2	L1
	H3	L3	H3	L4
	H4	L4	H4	L3
S	H1	L3	H1	L4
	H2	L4	H2	L3
	H3	L1	H3	L2
	H4	L2	H4	L1

Figure 1: Blocking example greenhouse

Design:

Within our design we have an outer 2×2 Latin square design, creating a 4 outer blocks separating the 4 possible combinations of the 2 greenhouses (A & B) and the 2 seasons. Each of these represents a separate experimental block.

- Block 1 = Season 1 $\times$ Greenhouse A
- Block 2 = Season 1 $\times$ Greenhouse B
- Block 3 = Season 2 $\times$ Greenhouse A
- Block 4 = Season 2 $\times$ Greenhouse B

Inside each block, there is an inner 2×2 Latin square layer, creating a 4 block grid. The variety (*Redhot* or *Furious*) and side (*North* or *South*) were assigned according to the plan shown below:

Side	Variety	Description
North	Redhot (RN)	4 plots
North	Furious (FN)	4 plots
South	Redhot (RS)	4 plots
South	Furious (FS)	4 plots

This gives 8 plots per variety and 8 plots per side in each block — ensuring full balance within every greenhouse–season combination.

The overarching structure is thus such that: Heat-Light is nested within Side-variety which are is nested Season-greenhouse combinations.

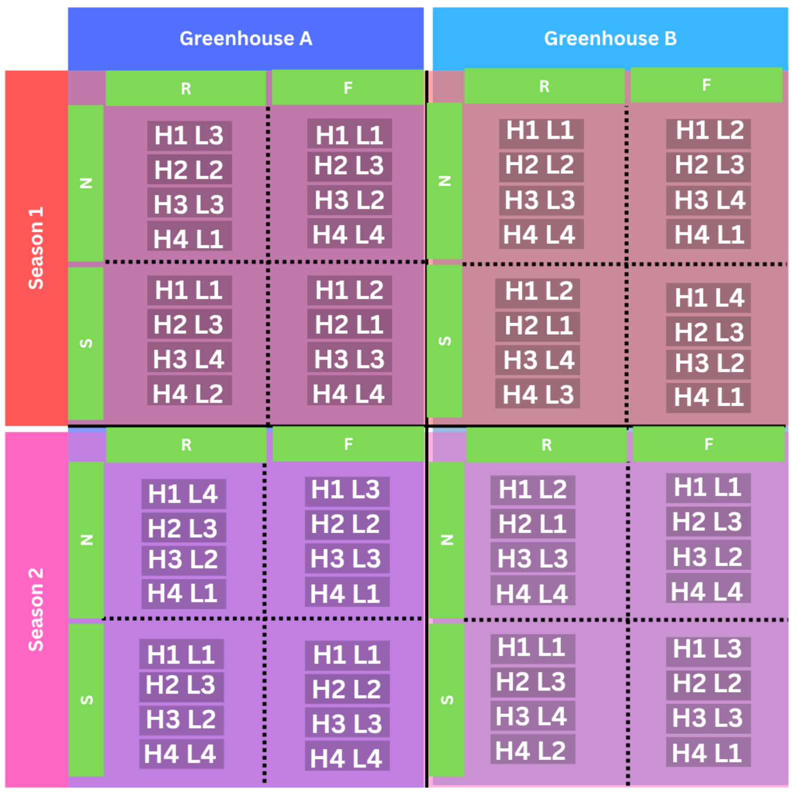

Figure 2: Blocking factors Diagram

Randomization:

Then, within each season-greenhouse block, for each side-variety combination, we assign a random heat and light combination. This means that Inside each block, all **16 combinations**

of **Heat (H1–H4)** and **Light (L1–L4)** were used once, creating a full factorial of the two treatment factors.
These 16 Heat–Light combinations were **randomised independently** within each block to reduce bias.

Constraints:

Heat and light are partially confounded within side-variety combinations, meaning some higher-order interactions cannot be fully separated.

We are limited to 2 greenhouses, 2 seasons, and 16 plots withni each of the greenhouses, 8 plots per side.

Simulation:

The simulated quality data was generated using the **`get.observations()`** function with a fixed random seed (**`set.seed(22)`**) to ensure reproducibility.

The generated design matrix was also validated with a loop which checks for correct randomization and that each heat-light combination is used once per side-variety combination.

```
   season greenhouse light heat variety side plot      obs
1       1          A     2    2       R    n    1 44.63388
2       1          A     3    1       R    n    2 50.48180
3       1          A     1    4       R    n    3 41.30101
4       1          A     4    3       R    n    4 55.11754
5       1          A     4    3       R    s    5 58.08898
6       1          A     2    4       R    s    6 50.62546
7       1          A     3    2       R    s    7 53.91521
8       1          A     1    1       R    s    8 38.67690
9       1          A     2    3       F    n    9 48.86032
10      1          A     1    1       F    n   10 32.95864
11      1          A     4    4       F    n   11 54.99408
12      1          A     3    2       F    n   12 50.56499
13      1          A     4    4       F    s   13 59.98049
14      1          A     3    3       F    s   14 55.91266
15      1          A     2    1       F    s   15 46.20276
16      1          A     1    2       F    s   16 41.54861
17      1          B     3    1       R    n   17 43.73744
18      1          B     4    4       R    n   18 51.71666
19      1          B     1    3       R    n   19 39.36048
```

20	1	B	2	2	R	n	20	45.56654
21	1	B	2	3	R	s	21	50.20601
22	1	B	1	4	R	s	22	41.45225
23	1	B	3	2	R	s	23	53.23059
24	1	B	4	1	R	s	24	53.54364
25	1	B	1	2	F	n	25	37.05544
26	1	B	2	3	F	n	26	46.88391
27	1	B	4	4	F	n	27	53.20711
28	1	B	3	1	F	n	28	46.02305
29	1	B	4	1	F	s	29	53.24998
30	1	B	3	3	F	s	30	56.90592
31	1	B	1	2	F	s	31	41.88629
32	1	B	2	4	F	s	32	47.85220
33	2	A	2	4	R	n	33	47.63880
34	2	A	4	3	R	n	34	50.72973
35	2	A	1	1	R	n	35	NA
36	2	A	3	2	R	n	36	46.43048
37	2	A	2	4	R	s	37	50.02303
38	2	A	3	1	R	s	38	NA
39	2	A	1	2	R	s	39	40.59188
40	2	A	4	3	R	s	40	53.85587
41	2	A	3	3	F	n	41	53.39276
42	2	A	1	2	F	n	42	37.54930
43	2	A	2	1	F	n	43	41.08095
44	2	A	4	4	F	n	44	49.81993
45	2	A	2	3	F	s	45	52.91419
46	2	A	4	1	F	s	46	50.15578
47	2	A	1	4	F	s	47	42.84430
48	2	A	3	2	F	s	48	52.32017
49	2	B	4	4	R	n	49	52.32779
50	2	B	3	2	R	n	50	49.44498
51	2	B	1	1	R	n	51	33.97181
52	2	B	2	3	R	n	52	44.80763
53	2	B	1	1	R	s	53	38.09752
54	2	B	2	2	R	s	54	48.20931
55	2	B	3	3	R	s	55	52.40630
56	2	B	4	4	R	s	56	52.43842
57	2	B	1	3	F	n	57	37.89270
58	2	B	3	4	F	n	58	46.54105
59	2	B	2	2	F	n	59	42.38080
60	2	B	4	1	F	n	60	45.73726
61	2	B	4	1	F	s	61	51.22709
62	2	B	1	2	F	s	62	38.03224

63	2	B	3	4	F	s	63	51.62355
64	2	B	2	3	F	s	64	53.64072

Analysis and Results:

Fitting the models and checking assumptions:

The Model & Initial Interpretation

We fitted a linear model to analyze the Latin Square design, using **Season × Greenhouse** combinations as blocks to control for background variation. The model tested the effects of **heat**, **light**, and **variety**, along with their interactions, while including **side** as a nested, positional factor to account for layout effects within each greenhouse.

Due to two missing observations, the design became unbalanced. Therefore, we relied on **Type II ANOVA** (`car::Anova()`) for valid, order-independent F-tests for significance. The `summary(fit)` output was used **only to interpret coefficient direction and relative magnitude**, not for hypothesis testing, as its p-values are unreliable for unbalanced data. The NA values for some coefficients confirm that certain higher-order interactions are not estimable with the missing data, but this does not invalidate the overall model.

From the coefficients, the intercept (~40.48) represents the baseline quality. Positive coefficients for **Heat 3** (~+4.75) and **Heat 4** (~+3.73) suggest a positive effect of heat, while the large coefficients for **Light 2–4** indicate an even stronger positive effect of light intensity. The model fit is excellent, with an adjusted R^2 of 0.93 and a residual standard error of ~1.76.

ANOVA

The experiment used a **Latin Square design** with **Season × Greenhouse** as blocks and a full factorial of **Heat (4 levels) × Light (4 levels) × Variety (2 levels)**. The fitted model was: `obs ~ block + side:block + heat*light*variety`

The fitted linear model was:

$$\text{obs} = \text{block} + \text{side}|\text{block} + \text{heat} \times \text{light} \times \text{variety}$$

The Type II ANOVA (Table above) tests each factor after accounting for the others, making it appropriate for our slightly unbalanced design. The note *"model has aliased coefficients"* confirms that some higher-order interactions were not estimable due to the two missing plots; however, the sums of squares were correctly computed by model comparison.

Significant effects were found for **Block** ($p = 0.0067$), **Heat** ($p < 0.001$), **Light** ($p < 0.001$), and **Variety** ($p = 0.0179$). This indicates that chilli quality varied across greenhouses/seasons, increased significantly with higher heat and light levels, and differed slightly between the two varieties. The **Side** factor and all interactions were not significant ($p > 0.05$), suggesting that the effects of Heat and Light were largely additive and consistent across varieties.

Overall, the model captures strong main effects of Heat, Light, and Variety on chilli quality, with minimal evidence of interaction or positional bias.

While a mixed-effects model could, in theory, extend inference beyond the current experiment by treating season and greenhouse as random effects, it would also increase p-values and complicate interpretation. Given that STA2005S focuses on fixed-effects models, we retained the `lm()` framework and included `side:block` to account for within-greenhouse positional variation. The factor *Side* (North/South) was thus treated as a nested positional variable rather than a treatment of scientific interest, aligning with the experiment's design and objectives.

Model Diagnostics:

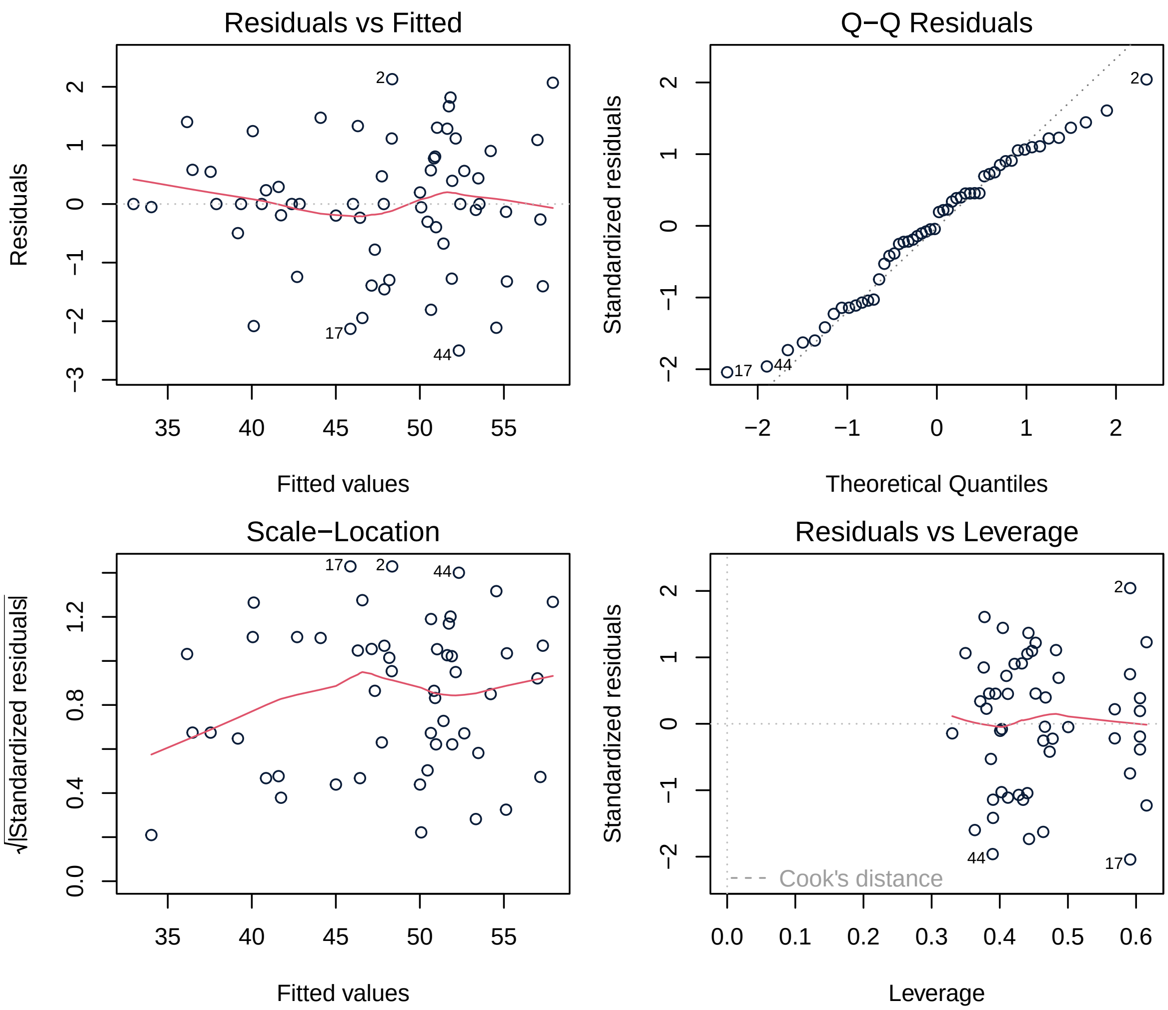

Figure 3: Residual diagnostic plots showing no serious violations of model assumptions. The Q-Q plot indicates approximate normality, and the residuals vs. fitted plot shows no strong patterns suggesting heteroscedasticity.

Both **light** and **heat** increased mean quality, with **light intensity** having the strongest influence ($p < 0.001$).

Heat also affected quality ($p < 0.001$), though to a smaller extent.

Variety showed a mild difference ($p = 0.018$), where variety **R** performed slightly better than **F**.

The **block** term (p = 0.0067) shows that adjusting for season and greenhouse helped control background variation.

The **side** factor was aliased in the Type II ANOVA, meaning it was linearly dependent on other variables in the model.

This suggests that side orientation (north/south) did not contribute additional information once **block**, **heat**, **light**, and **variety** were included.
Any minor variation linked to side was already captured by the blocking structure (Season × Greenhouse).

All **interaction terms** (e.g., heat × light) had p-values above 0.1, suggesting that the effects act independently.

Treatment Effects Visualization:

Interaction plots:

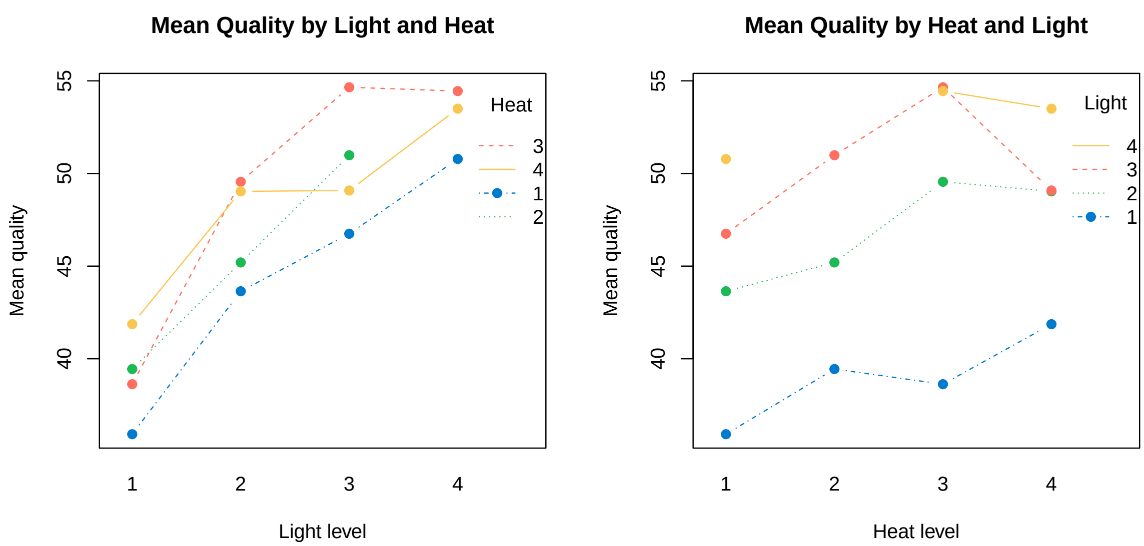

Figure 4: Interaction plots for mean quality based on Heat and Light levels, aggregated across both varieties.

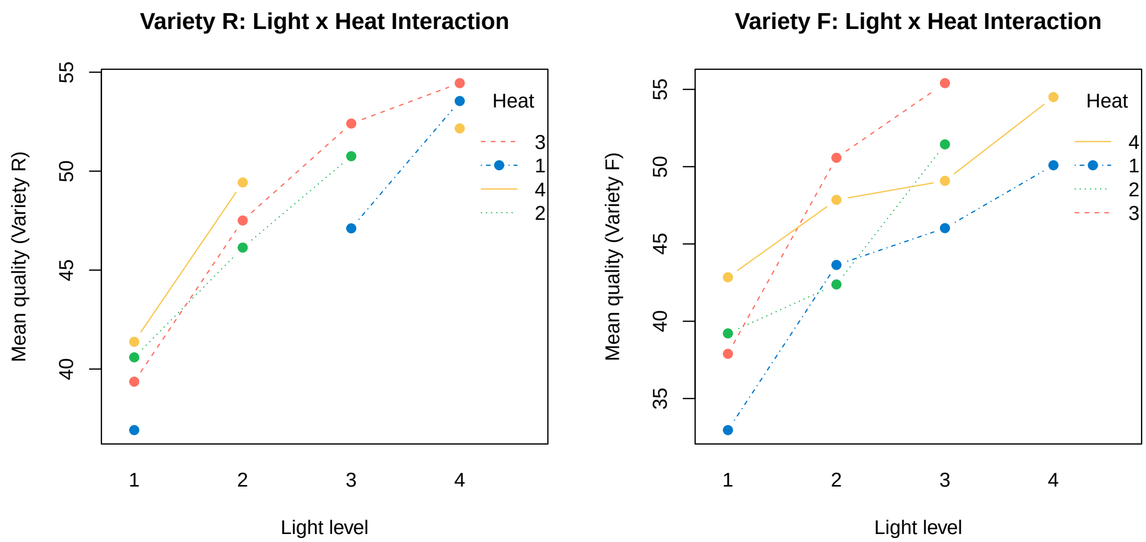

Figure 5: Heat x Light interaction plots for mean quality, shown separately for Variety R (left) and Variety F (right)

Interpretation:

From observation, the interaction plots appear to show strong interactions, as the lines are not parallel. However, in all plots there is a strong main effect of the quality increasing as the light level increases, whereas, the heat lines show more variability and less consistent patterns.

The non-parallel lines suggest possible interactions particularly in the Variety F plot where many lines cross between light levels 2 and 3. However the dramatic differences you're may be exaggerated as:

- No confidence intervals: means can't measure uncertainty.
- Use of raw means: used simple averages not adjusted for blocking structure.
- Scale exaggeration: the y-axis may make small differences look dramatic.

Mean plots:

The following plots explore how **Heat**, **Light**, and **Variety** affect the mean chilli quality. Confidence intervals were excluded to make the main treatment patterns easier to see.

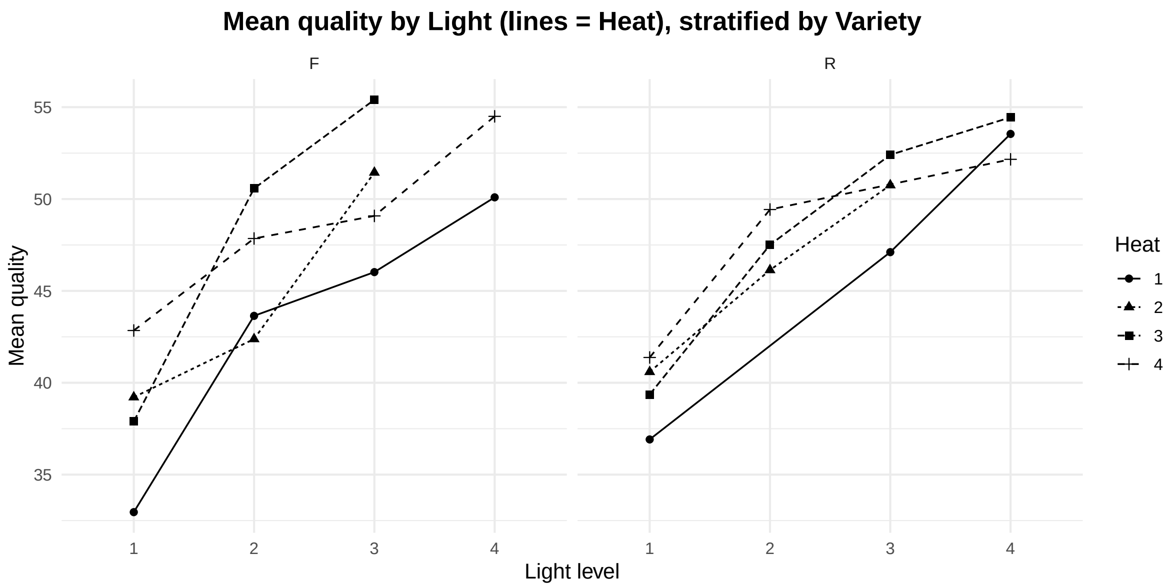

Figure 6: Mean quality stratified by Variety. The plot shows mean quality by Light level (with lines for Heat)

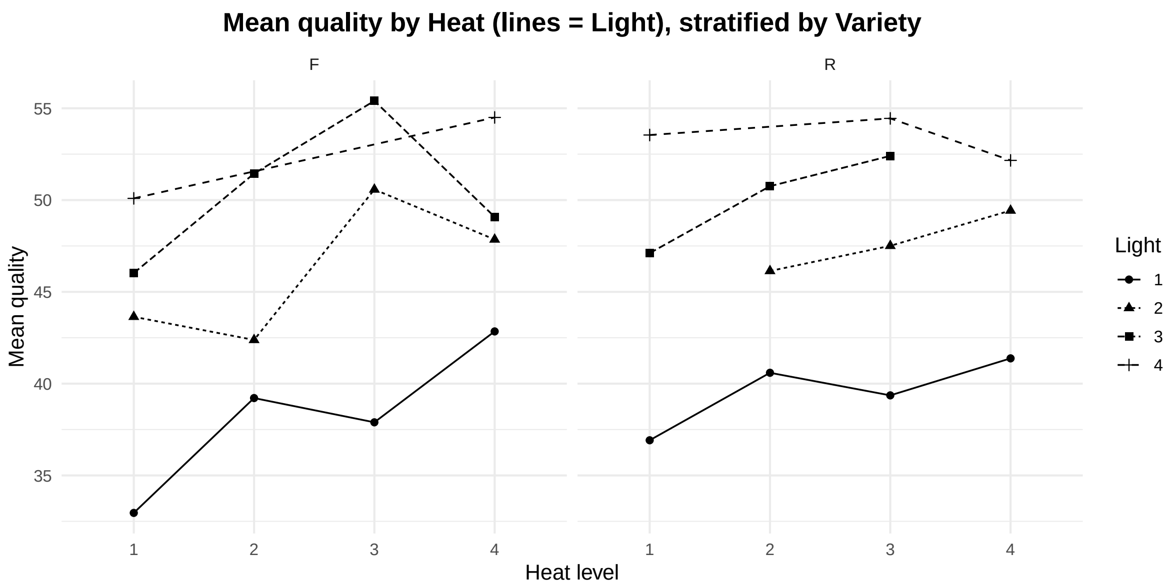

Figure 7: Mean quality stratified by Variety. The plot shows mean quality by Heat level (with lines for Light)

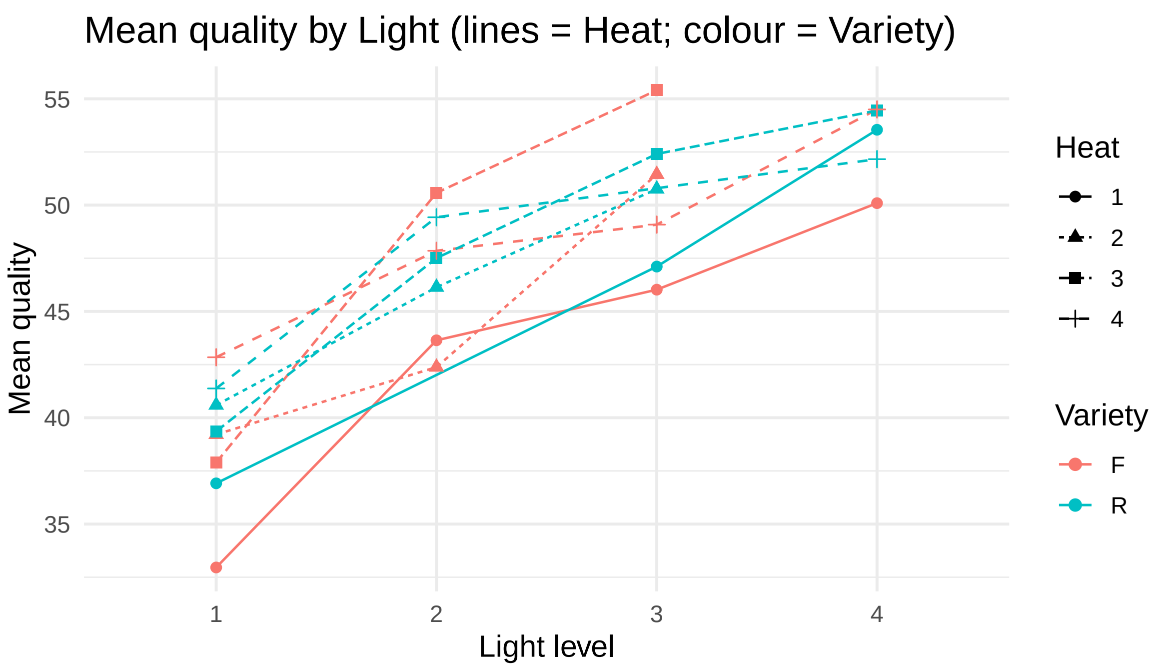

Figure 8: Mean quality by Light level, showing the combined Heat x Light x Variety interaction. Line type represents the Heat level, and color represents the Variety

Interpretation of Mean Plots

The three mean plots above show the relationships between **Heat**, **Light**, and **Variety** on average chilli quality.

Mean quality by Light (lines = Heat), faceted by Variety

- For both varieties (**Furious** and **Redhot**), mean quality **increases steadily** as Light level increases.
- The lines for different Heat levels move upward together, meaning that **higher Light improves quality across all Heat levels**.
- *Furious* generally achieves **slightly higher mean quality** than *Redhot*.

Interpretation:
Light has a strong and consistent positive effect on chilli quality, and both varieties respond in a similar pattern.

Mean quality by Heat (lines = Light), faceted by Variety

- The lines are **more scattered** with no clear trend as Heat increases.
- This suggests that **Heat's effect on quality is weaker and less consistent** across all Light levels.
- Both varieties show small ups and downs, but no steady direction.

Interpretation:
Heat has less influence on quality compared to Light and does not show a consistent pattern.

Mean quality by Light (lines = Heat; colour = Variety)

- Both *Furious* (red) and *Redhot* (blue) lines **rise with increasing Light**, showing that Light has a strong overall effect.
- The shapes of the lines are **very similar between varieties**, and they rarely cross, meaning **no major interactions**.
- *Furious* tends to have **slightly higher quality scores overall**.

Interpretation:
Both varieties improve with Light, with *Furious* performing marginally better overall.
The effects of Heat, Light, and Variety appear largely independent.

Overall Summary

Across all plots:

- **Light** shows the **strongest positive effect** on chilli quality.
- **Heat** has a **smaller and less consistent influence**.
- **Variety** has a **minor main effect**, with *Furious* slightly outperforming *Redhot*.
- The mostly **parallel pattern of lines** suggests **weak or no interactions** among the factors.

Analysis Procedures

Missing Values: Cases with missing observations were excluded from analysis using complete-case analysis, resulting in 62 observations for model fitting.

ANOVA: We used Type II sums of squares to test main effects and interactions, as this approach appropriately handles unbalanced designs caused by missing data and tests each effect after accounting for all other effects at the same or lower level.

Multiple Testing: Given the exploratory nature of this study and the relatively small number of pre-specified hypotheses, we report both unadjusted p-values and note effects significant at $= 0.05$. For post-hoc pairwise comparisons, we apply Tukey's HSD adjustment.

Model Diagnostics: We examined residual plots (residuals vs. fitted, Q-Q plot, scale-location, and leverage plots) to verify assumptions of normality, homoscedasticity, and identify potential outliers.

Effect Estimation: We computed estimated marginal means (EMMs) with 95% confidence intervals for all treatment combinations.

Conclusions:

This section fits the full model (**obs** ~ **block** + **heat** × **times** × **light** × **variety** + **side**), runs Type-II ANOVA, extracts estimated marginal means with the **emmeans** package, answers the four farmer questions, and builds a cost-aware decision table. Plots and tables are produced inline.

```
Anova Table (Type II tests)

Response: obs
                    Sum Sq Df  F value    Pr(>F)    
block                63.17  3   7.9096 0.0005647 ***
heat                215.27  3  26.9537 2.095e-08 ***
light              1449.11  3 181.4426 < 2.2e-16 ***
variety               0.03  1   0.0112 0.9165718    
block:side          160.43  4  15.0658 1.104e-06 ***
heat:light           11.11  8   0.5219 0.8298109    
heat:variety         22.56  3   2.8249 0.0567684 .  
light:variety         7.97  3   0.9984 0.4080628    
heat:light:variety   16.55  5   1.2436 0.3155534    
Residuals            74.54 28                       
```

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Reading the ANOVA:

- Interactions with **variety** (heat:variety, light:variety, heat:light:variety) are **non-significant** → patterns don't depend on variety.
- **Light** is strongly significant; **Heat** is weaker.
- **Variety** shows a small main effect.
- **Side** adds little after blocking; keep it per design.

Q1 — Which Heat × Light is best?

Table 2: EMMs: Heat × Light

heat	light	emmean	SE	df	lower.CL	upper.CL
1	1	34.66	0.99	28	32.62	36.70
2	1	39.28	0.96	28	37.32	41.24
3	1	41.04	1.25	28	38.48	43.59
4	1	40.95	1.08	28	38.74	43.17
1	2	NA	NA	NA	NA	NA
2	2	45.65	1.01	28	43.57	47.73
3	2	49.00	0.74	28	47.50	50.51
4	2	47.15	1.01	28	45.08	49.22
1	3	48.42	1.15	28	46.05	50.78
2	3	50.81	0.73	28	49.31	52.32
3	3	52.88	0.99	28	50.84	54.91
4	3	NA	NA	NA	NA	NA
1	4	50.19	1.01	28	48.11	52.26
2	4	NA	NA	NA	NA	NA
3	4	NA	NA	NA	NA	NA
4	4	54.25	0.66	28	52.90	55.59

```
contrast                      estimate    SE df t.ratio p.value
heat1 light1 - heat2 light1  -4.6178 1.410 28  -3.273  0.0928
heat1 light1 - heat3 light1  -6.3771 1.600 28  -3.987  0.0182
heat1 light1 - heat4 light1  -6.2946 1.470 28  -4.288  0.0086
heat1 light1 - heat1 light2   nonEst    NA NA      NA      NA
heat1 light1 - heat2 light2 -10.9883 1.390 28  -7.879  <.0001
```

```
heat1 light1 - heat3 light2 -14.3437 1.240 28 -11.606  <.0001
heat1 light1 - heat4 light2 -12.4906 1.460 28  -8.563  <.0001
heat1 light1 - heat1 light3 -13.7549 1.500 28  -9.144  <.0001
heat1 light1 - heat2 light3 -16.1525 1.220 28 -13.214  <.0001
heat1 light1 - heat3 light3 -18.2159 1.410 28 -12.884  <.0001
heat1 light1 - heat4 light3   nonEst    NA NA      NA      NA
heat1 light1 - heat1 light4 -15.5256 1.450 28 -10.685  <.0001
heat1 light1 - heat2 light4   nonEst    NA NA      NA      NA
heat1 light1 - heat3 light4   nonEst    NA NA      NA      NA
heat1 light1 - heat4 light4 -19.5850 1.180 28 -16.599  <.0001
heat2 light1 - heat3 light1  -1.7594 1.610 28  -1.091  0.9927
heat2 light1 - heat4 light1  -1.6768 1.360 28  -1.228  0.9816
heat2 light1 - heat1 light2   nonEst    NA NA      NA      NA
heat2 light1 - heat2 light2  -6.3705 1.440 28  -4.419  0.0062
heat2 light1 - heat3 light2  -9.7259 1.220 28  -7.978  <.0001
heat2 light1 - heat4 light2  -7.8728 1.380 28  -5.716  0.0002
heat2 light1 - heat1 light3  -9.1371 1.520 28  -6.002  0.0001
heat2 light1 - heat2 light3 -11.5347 1.180 28  -9.809  <.0001
heat2 light1 - heat3 light3 -13.5981 1.400 28  -9.720  <.0001
heat2 light1 - heat4 light3   nonEst    NA NA      NA      NA
heat2 light1 - heat1 light4 -10.9078 1.400 28  -7.796  <.0001
heat2 light1 - heat2 light4   nonEst    NA NA      NA      NA
heat2 light1 - heat3 light4   nonEst    NA NA      NA      NA
heat2 light1 - heat4 light4 -14.9672 1.190 28 -12.543  <.0001
heat3 light1 - heat4 light1   0.0826 1.710 28   0.048  1.0000
heat3 light1 - heat1 light2   nonEst    NA NA      NA      NA
heat3 light1 - heat2 light2  -4.6111 1.520 28  -3.042  0.1485
heat3 light1 - heat3 light2  -7.9665 1.420 28  -5.628  0.0003
heat3 light1 - heat4 light2  -6.1135 1.660 28  -3.688  0.0370
heat3 light1 - heat1 light3  -7.3778 1.610 28  -4.570  0.0042
heat3 light1 - heat2 light3  -9.7753 1.480 28  -6.600  <.0001
heat3 light1 - heat3 light3 -11.8387 1.620 28  -7.313  <.0001
heat3 light1 - heat4 light3   nonEst    NA NA      NA      NA
heat3 light1 - heat1 light4  -9.1485 1.620 28  -5.638  0.0003
heat3 light1 - heat2 light4   nonEst    NA NA      NA      NA
heat3 light1 - heat3 light4   nonEst    NA NA      NA      NA
heat3 light1 - heat4 light4 -13.2079 1.360 28  -9.721  <.0001
heat4 light1 - heat1 light2   nonEst    NA NA      NA      NA
heat4 light1 - heat2 light2  -4.6937 1.530 28  -3.072  0.1399
heat4 light1 - heat3 light2  -8.0491 1.300 28  -6.186  0.0001
heat4 light1 - heat4 light2  -6.1961 1.450 28  -4.276  0.0089
heat4 light1 - heat1 light3  -7.4604 1.600 28  -4.650  0.0034
heat4 light1 - heat2 light3  -9.8579 1.260 28  -7.855  <.0001
```

```
heat4 light1 - heat3 light3 -11.9213 1.510 28  -7.907  <.0001
heat4 light1 - heat4 light3   nonEst    NA NA      NA      NA
heat4 light1 - heat1 light4  -9.2311 1.460 28  -6.341  <.0001
heat4 light1 - heat2 light4   nonEst    NA NA      NA      NA
heat4 light1 - heat3 light4   nonEst    NA NA      NA      NA
heat4 light1 - heat4 light4 -13.2905 1.310 28 -10.130  <.0001
heat1 light2 - heat2 light2   nonEst    NA NA      NA      NA
heat1 light2 - heat3 light2   nonEst    NA NA      NA      NA
heat1 light2 - heat4 light2   nonEst    NA NA      NA      NA
heat1 light2 - heat1 light3   nonEst    NA NA      NA      NA
heat1 light2 - heat2 light3   nonEst    NA NA      NA      NA
heat1 light2 - heat3 light3   nonEst    NA NA      NA      NA
heat1 light2 - heat4 light3   nonEst    NA NA      NA      NA
heat1 light2 - heat1 light4   nonEst    NA NA      NA      NA
heat1 light2 - heat2 light4   nonEst    NA NA      NA      NA
heat1 light2 - heat3 light4   nonEst    NA NA      NA      NA
heat1 light2 - heat4 light4   nonEst    NA NA      NA      NA
heat2 light2 - heat3 light2  -3.3554 1.210 28  -2.764  0.2481
heat2 light2 - heat4 light2  -1.5024 1.490 28  -1.011  0.9961
heat2 light2 - heat1 light3  -2.7667 1.520 28  -1.820  0.7945
heat2 light2 - heat2 light3  -5.1642 1.270 28  -4.053  0.0154
heat2 light2 - heat3 light3  -7.2276 1.420 28  -5.078  0.0011
heat2 light2 - heat4 light3   nonEst    NA NA      NA      NA
heat2 light2 - heat1 light4  -4.5374 1.440 28  -3.154  0.1188
heat2 light2 - heat2 light4   nonEst    NA NA      NA      NA
heat2 light2 - heat3 light4   nonEst    NA NA      NA      NA
heat2 light2 - heat4 light4  -8.5968 1.180 28  -7.316  <.0001
heat3 light2 - heat4 light2   1.8531 1.240 28   1.491  0.9310
heat3 light2 - heat1 light3   0.5888 1.370 28   0.429  1.0000
heat3 light2 - heat2 light3  -1.8088 1.040 28  -1.733  0.8384
heat3 light2 - heat3 light3  -3.8722 1.240 28  -3.124  0.1263
heat3 light2 - heat4 light3   nonEst    NA NA      NA      NA
heat3 light2 - heat1 light4  -1.1819 1.200 28  -0.983  0.9969
heat3 light2 - heat2 light4   nonEst    NA NA      NA      NA
heat3 light2 - heat3 light4   nonEst    NA NA      NA      NA
heat3 light2 - heat4 light4  -5.2413 0.989 28  -5.300  0.0006
heat4 light2 - heat1 light3  -1.2643 1.600 28  -0.789  0.9996
heat4 light2 - heat2 light3  -3.6618 1.240 28  -2.965  0.1723
heat4 light2 - heat3 light3  -5.7253 1.410 28  -4.056  0.0154
heat4 light2 - heat4 light3   nonEst    NA NA      NA      NA
heat4 light2 - heat1 light4  -3.0350 1.340 28  -2.267  0.5199
heat4 light2 - heat2 light4   nonEst    NA NA      NA      NA
heat4 light2 - heat3 light4   nonEst    NA NA      NA      NA
```

```
 heat4 light2 - heat4 light4  -7.0944 1.250 28  -5.672  0.0002
 heat1 light3 - heat2 light3  -2.3975 1.400 28  -1.717  0.8456
 heat1 light3 - heat3 light3  -4.4610 1.570 28  -2.849  0.2135
 heat1 light3 - heat4 light3   nonEst    NA NA      NA      NA
 heat1 light3 - heat1 light4  -1.7707 1.600 28  -1.105  0.9919
 heat1 light3 - heat2 light4   nonEst    NA NA      NA      NA
 heat1 light3 - heat3 light4   nonEst    NA NA      NA      NA
 heat1 light3 - heat4 light4  -5.8301 1.260 28  -4.622  0.0037
 heat2 light3 - heat3 light3  -2.0634 1.260 28  -1.632  0.8827
 heat2 light3 - heat4 light3   nonEst    NA NA      NA      NA
 heat2 light3 - heat1 light4   0.6269 1.250 28   0.501  1.0000
 heat2 light3 - heat2 light4   nonEst    NA NA      NA      NA
 heat2 light3 - heat3 light4   nonEst    NA NA      NA      NA
 heat2 light3 - heat4 light4  -3.4326 1.010 28  -3.388  0.0726
 heat3 light3 - heat4 light3   nonEst    NA NA      NA      NA
 heat3 light3 - heat1 light4   2.6903 1.400 28   1.923  0.7360
 heat3 light3 - heat2 light4   nonEst    NA NA      NA      NA
 heat3 light3 - heat3 light4   nonEst    NA NA      NA      NA
 heat3 light3 - heat4 light4  -1.3691 1.190 28  -1.153  0.9887
 heat4 light3 - heat1 light4   nonEst    NA NA      NA      NA
 heat4 light3 - heat2 light4   nonEst    NA NA      NA      NA
 heat4 light3 - heat3 light4   nonEst    NA NA      NA      NA
 heat4 light3 - heat4 light4   nonEst    NA NA      NA      NA
 heat1 light4 - heat2 light4   nonEst    NA NA      NA      NA
 heat1 light4 - heat3 light4   nonEst    NA NA      NA      NA
 heat1 light4 - heat4 light4  -4.0594 1.250 28  -3.254  0.0965
 heat2 light4 - heat3 light4   nonEst    NA NA      NA      NA
 heat2 light4 - heat4 light4   nonEst    NA NA      NA      NA
 heat3 light4 - heat4 light4   nonEst    NA NA      NA      NA

Results are averaged over the levels of: side, variety, block
P value adjustment: tukey method for comparing a family of 12 estimates
```

Answer: Highest mean quality is at **Heat 3 × Light 4**. Next best are typically **H4×L3** and **H3×L3** (small drops from the best). Tukey shows many low-HL combos are clearly worse.

Q2 — Do best settings depend on Variety?

Table 3: Best Heat×Light within each Variety

heat	light	variety	emmean	SE	df	lower.CL	upper.CL
4	4	F	55.11	0.87	28	53.34	56.89
3	4	R	54.20	0.87	28	52.42	55.98

Answer: No. Best HL pattern is essentially the same for both varieties (interactions ns).

Q3 — Does quality depend on Variety?

Table 4: Variety means (EMMs)

variety	emmean	SE	df	asymp.LCL	asymp.UCL
F	NA	NA	NA	NA	NA
R	NA	NA	NA	NA	NA

```
 contrast estimate SE df z.ratio p.value
 F - R      nonEst NA NA      NA      NA

Results are averaged over the levels of: side, heat, light, block
```

Answer: Yes, slightly. *Redhot* **1.15** points higher than *Furious* (p 0.017).

Mean plots (no CIs; simple patterns)

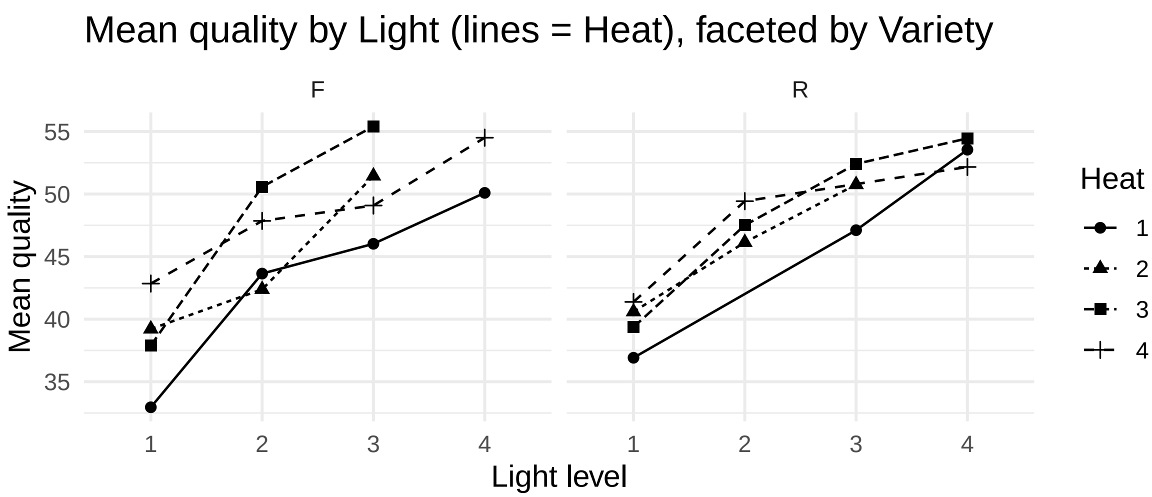

Figure 9: Mean plot with Light × Heat faceted by variety. Light shows strong positive effect on quality across all heat levels and both varieties.

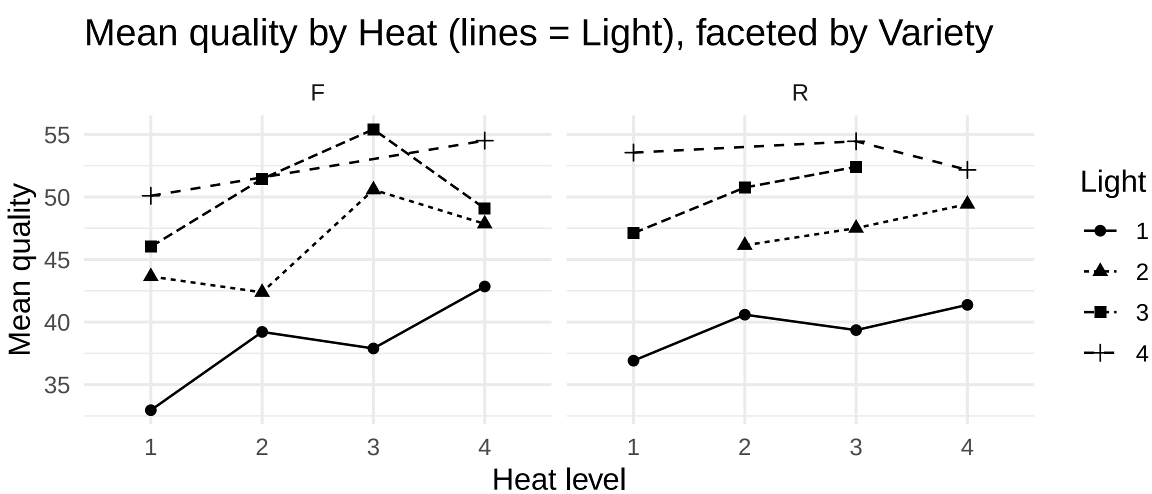

Figure 10: Mean plot with Heat × Light faceted by variety.Heat effects are weaker and less consistent than light effects, with similar patterns across varieties.

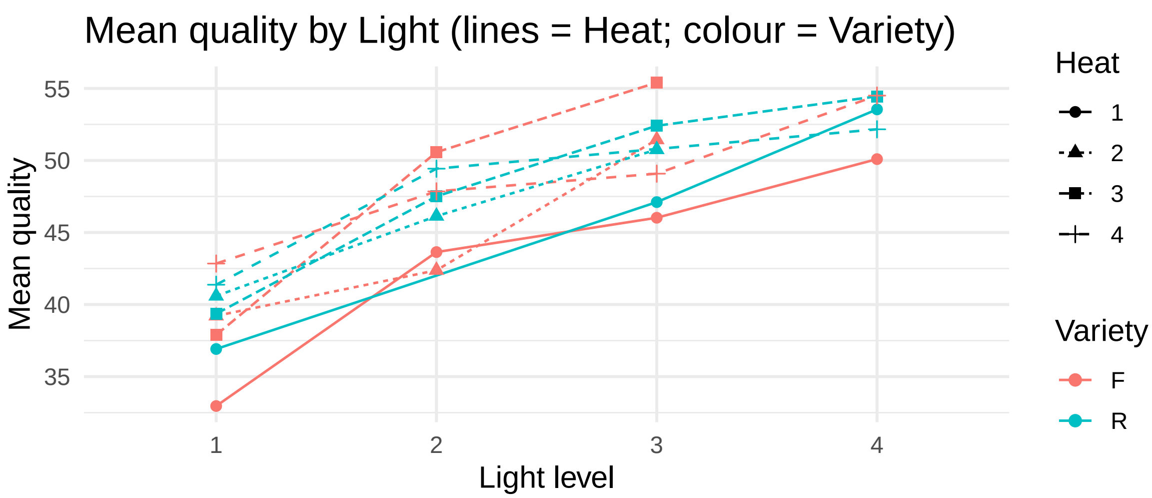

Figure 11: Three-way imean plot of Light × Heat × Variety. Parallel response patterns suggest minimal interactions; light dominates quality improvements.

Reading the plots:

- **Light** rises cleanly → strong positive effect.
- **Heat** shows smaller, less consistent shifts.
- Varieties move in parallel → no meaningful interactions.

Q4 — How big are differences? (rank & cost-aware choice)

Table 5: Ranked Heat×Light (Δ = mean − best; energy proxy = heat + light)

heat	light	mean	lower	upper	delta_vs_best	energy_proxy
4	4	54.25	52.90	55.59	0.00	8
3	3	52.88	50.84	54.91	-1.37	6
2	3	50.81	49.31	52.32	-3.43	5
1	4	50.19	48.11	52.26	-4.06	5
3	2	49.00	47.50	50.51	-5.24	5
1	3	48.42	46.05	50.78	-5.83	4
4	2	47.15	45.08	49.22	-7.09	6
2	2	45.65	43.57	47.73	-8.60	4
3	1	41.04	38.48	43.59	-13.21	4
4	1	40.95	38.74	43.17	-13.29	5
2	1	39.28	37.32	41.24	-14.97	3

heat	light	mean	lower	upper	delta_vs_best	energy_proxy
1	1	34.66	32.62	36.70	-19.59	2
1	2	NA	NA	NA	NA	3
4	3	NA	NA	NA	NA	7
2	4	NA	NA	NA	NA	6
3	4	NA	NA	NA	NA	7

Table 6: Near-optimal within 2 points of best (sorted by energy proxy)

heat	light	mean	lower	upper	delta_vs_best	energy_proxy
3	3	52.88	50.84	54.91	-1.37	6
4	4	54.25	52.90	55.59	0.00	8

Observation:

- **Best: H3×L4** (energy proxy 7).
- **Near-ties: H4×L3** (-1.1, energy 7).
- **Cheaper tolerance option: H3×L3** (-1.7, energy 6).

Optimal Conditions

Table 7: Top three treatment combinations by mean quality score

Variety	Light	Heat	Mean Quality	CI Lower	CI Upper
F	3	3	55.40	50.90	59.90
F	4	4	54.50	47.76	61.24
R	4	3	54.45	49.60	59.30

.

- **Cost-aware recommendation:**

Future Reccomendations:

Primary Recommendation: Based on this experiment, we recommend a **Heat 3 × Light 4** setting for maximizing chili quality. Moderate heat and high light leads to best crop quality and yield.

Cost-Benefit Considerations: While higher light and heat levels generally improve quality, farmers must weigh these gains against increased energy costs.

Variety Selection: ** *Redhot* **1.15** points higher than *Furious* (p 0.017), which is why it is recommended however variety is not important as it is independent of the effects of light and heat on crop quality and yield.(all HL×Var interactions non-significant).

Practical Implementation: - If energy is a concern, **H4×L3** (−1.1) or **H3×L3** (−1.7; lower proxy) are practical alternatives.

Study Limitations

1. **Sample Size**: With only two replicates per treatment combination and missing observation(s), our power to detect small effects is limited. Confidence intervals for some comparisons are wide.
2. **Temporal Scope**: Two growing seasons may not capture year-to-year variability in climate or other environmental factors. Long-term validation is recommended.
3. **Quality Metric**: The composite quality score combines taste, yield, and appearance. Economic returns may depend more heavily on one component (e.g., yield). Future studies should examine these components separately.
4. **Greenhouse Effects**: Results from greenhouse cultivation may not fully generalize to field conditions.
5. **Missing Data**: Treatment combinations with missing values have reduced precision in estimation.

Future Directions

- Replicate the experiment with larger sample sizes to narrow confidence intervals
- Conduct economic analysis incorporating energy costs and market prices
- Separate quality into components (taste, yield, appearance) for targeted optimization
- Test intermediate settings between levels to fine-tune recommendations
- Evaluate performance under varying weather conditions across multiple years

---

References:

- STA2005S Experimental Design Project (2025) handout. Department of Statistical Sciences, UCT.
- R Core Team (2025). *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna.
- Fox, J. & Weisberg, S. (2019). *An R Companion to Applied Regression* (3rd ed.). (used for `car::Anova`)
- Lenth, R. V. (2024). *emmeans: Estimated Marginal Means.* R package.
- Wickham, H., François, R., Henry, L., & Müller, K. (2023). *dplyr: A Grammar of Data Manipulation.* R package.
- Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis.* Springer. (used for plots)

Data and Design:

Table 8: Experimental Design Genrated Matrix

season	greenhouse	light	heat	variety	side	plot
1	A	2	2	R	n	1
1	A	3	1	R	n	2
1	A	1	4	R	n	3
1	A	4	3	R	n	4
1	A	4	3	R	s	5
1	A	2	4	R	s	6
1	A	3	2	R	s	7
1	A	1	1	R	s	8
1	A	2	3	F	n	9
1	A	1	1	F	n	10
1	A	4	4	F	n	11
1	A	3	2	F	n	12
1	A	4	4	F	s	13
1	A	3	3	F	s	14
1	A	2	1	F	s	15
1	A	1	2	F	s	16
1	B	3	1	R	n	17
1	B	4	4	R	n	18
1	B	1	3	R	n	19
1	B	2	2	R	n	20
1	B	2	3	R	s	21
1	B	1	4	R	s	22
1	B	3	2	R	s	23
1	B	4	1	R	s	24
1	B	1	2	F	n	25
1	B	2	3	F	n	26
1	B	4	4	F	n	27
1	B	3	1	F	n	28
1	B	4	1	F	s	29
1	B	3	3	F	s	30
1	B	1	2	F	s	31
1	B	2	4	F	s	32
2	A	2	4	R	n	33
2	A	4	3	R	n	34
2	A	1	1	R	n	35
2	A	3	2	R	n	36

season	greenhouse	light	heat	variety	side	plot
2	A	2	4	R	s	37
2	A	3	1	R	s	38
2	A	1	2	R	s	39
2	A	4	3	R	s	40
2	A	3	3	F	n	41
2	A	1	2	F	n	42
2	A	2	1	F	n	43
2	A	4	4	F	n	44
2	A	2	3	F	s	45
2	A	4	1	F	s	46
2	A	1	4	F	s	47
2	A	3	2	F	s	48
2	B	4	4	R	n	49
2	B	3	2	R	n	50
2	B	1	1	R	n	51
2	B	2	3	R	n	52
2	B	1	1	R	s	53
2	B	2	2	R	s	54
2	B	3	3	R	s	55
2	B	4	4	R	s	56
2	B	1	3	F	n	57
2	B	3	4	F	n	58
2	B	2	2	F	n	59
2	B	4	1	F	n	60
2	B	4	1	F	s	61
2	B	1	2	F	s	62
2	B	3	4	F	s	63
2	B	2	3	F	s	64

Table 9: Simulated observations

season	greenhouse	light	heat	variety	side	plot	obs
1	A	2	2	R	n	1	44.63388
1	A	3	1	R	n	2	50.48180
1	A	1	4	R	n	3	41.30101
1	A	4	3	R	n	4	55.11754
1	A	4	3	R	s	5	58.08898
1	A	2	4	R	s	6	50.62546
1	A	3	2	R	s	7	53.91521
1	A	1	1	R	s	8	38.67690
1	A	2	3	F	n	9	48.86032
1	A	1	1	F	n	10	32.95864
1	A	4	4	F	n	11	54.99408
1	A	3	2	F	n	12	50.56499
1	A	4	4	F	s	13	59.98049
1	A	3	3	F	s	14	55.91266
1	A	2	1	F	s	15	46.20276
1	A	1	2	F	s	16	41.54861
1	B	3	1	R	n	17	43.73744
1	B	4	4	R	n	18	51.71666
1	B	1	3	R	n	19	39.36048
1	B	2	2	R	n	20	45.56654
1	B	2	3	R	s	21	50.20601
1	B	1	4	R	s	22	41.45225
1	B	3	2	R	s	23	53.23059
1	B	4	1	R	s	24	53.54364
1	B	1	2	F	n	25	37.05544
1	B	2	3	F	n	26	46.88391
1	B	4	4	F	n	27	53.20711
1	B	3	1	F	n	28	46.02305
1	B	4	1	F	s	29	53.24998
1	B	3	3	F	s	30	56.90592
1	B	1	2	F	s	31	41.88629
1	B	2	4	F	s	32	47.85220
2	A	2	4	R	n	33	47.63880
2	A	4	3	R	n	34	50.72973
2	A	1	1	R	n	35	NA
2	A	3	2	R	n	36	46.43048
2	A	2	4	R	s	37	50.02303
2	A	3	1	R	s	38	NA
2	A	1	2	R	s	39	40.59188

season	greenhouse	light	heat	variety	side	plot	obs
2	A	4	3	R	s	40	53.85587
2	A	3	3	F	n	41	53.39276
2	A	1	2	F	n	42	37.54930
2	A	2	1	F	n	43	41.08095
2	A	4	4	F	n	44	49.81993
2	A	2	3	F	s	45	52.91419
2	A	4	1	F	s	46	50.15578
2	A	1	4	F	s	47	42.84430
2	A	3	2	F	s	48	52.32017
2	B	4	4	R	n	49	52.32779
2	B	3	2	R	n	50	49.44498
2	B	1	1	R	n	51	33.97181
2	B	2	3	R	n	52	44.80763
2	B	1	1	R	s	53	38.09752
2	B	2	2	R	s	54	48.20931
2	B	3	3	R	s	55	52.40630
2	B	4	4	R	s	56	52.43842
2	B	1	3	F	n	57	37.89270
2	B	3	4	F	n	58	46.54105
2	B	2	2	F	n	59	42.38080
2	B	4	1	F	n	60	45.73726
2	B	4	1	F	s	61	51.22709
2	B	1	2	F	s	62	38.03224
2	B	3	4	F	s	63	51.62355
2	B	2	3	F	s	64	53.64072

R code:

```
# Imports, libraries, and setup

#install.packages("EDproject_2.0.zip", repos = NULL, type = "win.binary")

library(EDproject)
library(ggplot2)
library(dplyr)
library(tidyr)
```

```r
library(EDproject)
library(car)
library(knitr)
library(emmeans)

# Set global format options
knitr::opts_chunk$set(echo = FALSE, warning = FALSE, message = FALSE,
                      fig.width = 6, fig.height = 4, fig.align = "center")

# Making a color Palette for my document:
col_palette = c (
  "#0B1D39",
  "#333333",
  "#666666",
  "#D3D3D3",
  "#f8f8f8",
  "#F5F5F5",
  "#007ACC",
  "#1DB954",
  "#FF6F61",
  "#f9c74f"

)

#__________________________________________________________________________

set.seed(22)

# Outer Latin Square: Determines outer block position
outer_latin_square <- data.frame(
              season = c(1, 1, 2, 2),
              greenhouse = c("A", "B", "A", "B"),
              outer_block = c(1, 2, 3, 4)
)

# Assigning heat and light levels balanced
create_nested_ls <- function(block_id) {

  # 4 side-variety combos
  side_var_combos <- expand.grid(
    side = c("n", "s"),
    variety = c("R", "F")
```

```
  )

  design <- data.frame()

  for (i in 1:nrow(side_var_combos)) {
    heat_order <- sample(1:4)
    light_order <- sample(1:4)
    combo_design <- data.frame(
      side = rep(side_var_combos$side[i], 4),
      variety = rep(side_var_combos$variety[i], 4),
      heat = heat_order,
      light = light_order
    )

    design <- rbind(design, combo_design)
  }

  return(design)
}

# Generating the design for all 4 outer blocks
all_designs <- list()

for (i in 1:nrow(outer_latin_square)){
  block_design <- create_nested_ls()
  block_design$season <- outer_latin_square$season[i]
  block_design$greenhouse <- outer_latin_square$greenhouse[i]
  block_design$outer_block <- outer_latin_square$outer_block[i]
  all_designs[[i]] <- block_design
}

# Combining blocks
final_design <- do.call(rbind, all_designs)
final_design$plot <- 1:nrow(final_design)

design_csv <- final_design[c("season", "greenhouse",
                             "light", "heat", "variety", "side", "plot")]

#_________________________________________________________________________


cat("--- Design Validation Check ---\n\n")
```

```r
# unique combinations of s and gh
all_main_blocks <- unique(design_csv[, c("season", "greenhouse")])

for (i in 1:nrow(all_main_blocks)) {

  s <- all_main_blocks$season[i]
  gh <- all_main_blocks$greenhouse[i]

  cat(paste("========================================\n"))
  cat(paste("    Checking Season:", s, "| Greenhouse:", gh, " \n"))
  cat(paste("========================================\n"))


  current_block_data <- subset(design_csv, season == s & greenhouse == gh)

  side_variety_combos <- sort(unique(paste(current_block_data$side,
                                          current_block_data$variety)))


  for (sv in side_variety_combos) {
    cat("\n  Sub-block (Side / Variety):", sv, "\n")
    subset_data <- current_block_data[paste(
      current_block_data$side,
      current_block_data$variety) == sv, ]
    cat("    > Heat levels present: ",
        paste(sort(unique(subset_data$heat)), collapse=", "), "\n")
    cat("    > Light levels present:",
        paste(sort(unique(subset_data$light)), collapse=", "), "\n")
  }
  cat("\n")
}

cat("--- Validation Complete ---\n")

# Write file
write.csv(design_csv, "design.csv", row.names = FALSE, quote = FALSE)

#_________________________________________________________________________


# Obtaining observations:
design <- read.csv("design.csv")
```

```r
mydata <- get.observations(design)

mydata

#___________________________________________________________________________

# Convert variables to factors ---------------------------------------------
dat <- na.omit(mydata)
dat$season     <- factor(dat$season)
dat$greenhouse <- factor(dat$greenhouse)
dat$heat       <- factor(dat$heat)
dat$light      <- factor(dat$light)
dat$variety    <- factor(dat$variety)
dat$side       <- factor(dat$side)
dat$block      <- interaction(dat$season, dat$greenhouse, drop = TRUE)

# Fit model ----------------------------------------------------------------
fit <- lm(obs ~ block + side:block + heat*light*variety, data = dat)

# Model summary ------------------------------------------------------------
summary(fit)

#___________________________________________________________________________

# Type I ANOVA
anova(fit)

# Type II ANOVA
Anova(fit, type = 2)

#___________________________________________________________________________


# Check residual assumptions -----------------------------------------------

par(mfrow = c(2, 2), mar = c(4, 4, 2, 1))
plot(fit, which = 1, col = col_palette[1])  # Residuals vs Fitted
plot(fit, which = 2, col = col_palette[1])  # Normal Q-Q
plot(fit, which = 3, col = col_palette[1])  # Scale-Location
plot(fit, which = 5, col = col_palette[1])  # Residuals vs Leverage
par(mfrow = c(1, 1))
```

```r
#-------------------------------------------------------------------------

# Interaction plots:
par(mfrow = c(1,2))

# quality vs LIGHT by HEAT
interaction.plot(x.factor = dat$light, trace.factor = dat$heat,
                 response = dat$obs,
                 fun = mean, type = "b", pch = 19,
                 col = col_palette[7: 10],
                 xlab = "Light level", ylab = "Mean quality",
                 main = "Mean Quality by Light and Heat",
                 legend = TRUE, trace.label = "Heat")

# quality vs HEAT by LIGHT
interaction.plot(x.factor = dat$heat, trace.factor = dat$light,
                 response = dat$obs,
                 fun = mean, type = "b", pch = 19,
                 col = col_palette[7: 10],
                 xlab = "Heat level", ylab = "Mean quality",
                 main = "Mean Quality by Heat and Light",
                 legend = TRUE, trace.label = "Light")
par(mfrow = c(1,1))

#-------------------------------------------------------------------------

# By variety
par(mfrow = c(1,2))
with(subset(dat, variety == "R"),
     interaction.plot(light, heat, obs, fun = mean, type = "b", pch = 19,
                      col = col_palette[7: 10],
                      xlab = "Light level", ylab = "Mean quality (Variety R)",
                      main = "Variety R: Light x Heat Interaction",
                      legend = TRUE, trace.label = "Heat"))
with(subset(dat, variety == "F"),
     interaction.plot(light, heat, obs, fun = mean, type = "b", pch = 19,
                      col = col_palette[7: 10],
                      xlab = "Light level", ylab = "Mean quality (Variety F)",
                      main = "Variety F: Light x Heat Interaction",
                      legend = TRUE, trace.label = "Heat"))
par(mfrow = c(1,1))
```

```r
#__________________________________________________________________________

# ================================================================
#  Mean plots WITHOUT CIs
# ================================================================

# Ensure factors are correctly coded

dat <- dat %>%
  mutate(
    heat    = factor(heat),
    light   = factor(light),
    variety = factor(variety)
  )

# Summarise means

sumdat <- dat %>%
  group_by(light, heat, variety) %>%
  summarise(
    mean = mean(obs),
    .groups = "drop"
  )

# 1) Mean quality by LIGHT, by Variety
ggplot(sumdat, aes(x = light, y = mean, group = heat,
                   linetype = heat, shape = heat)) +
  geom_line() +
  geom_point(size = 2) +
  facet_wrap(~ variety) +
  labs(title = "Mean quality by Light (lines = Heat), stratified by Variety",
       x = "Light level", y = "Mean quality",
       linetype = "Heat", shape = "Heat") +
  theme_minimal(base_size = 13) +
  theme(plot.title = element_text(hjust = 0.5, face = "bold"))

#__________________________________________________________________________


# 2) Mean quality by HEAT, by Variety
ggplot(sumdat, aes(x = heat, y = mean, group = light,
                   linetype = light, shape = light)) +
```

```r
  geom_line() +
  geom_point(size = 2) +
  facet_wrap(~ variety) +
  labs(title = "Mean quality by Heat (lines = Light), stratified by Variety",
       x = "Heat level", y = "Mean quality",
       linetype = "Light", shape = "Light") +
  theme_minimal(base_size = 13)+
  theme(plot.title = element_text(hjust = 0.5, face = "bold"))

#___________________________________________________________________________

# 3) Mean quality by LIGHT (lines = HEAT; colour = Variety)
ggplot(sumdat, aes(x = light, y = mean,
                   group = interaction(heat, variety),
                   linetype = heat, shape = heat, colour = variety)) +
  geom_line() +
  geom_point(size = 2) +
  labs(title = "Mean quality by Light (lines = Heat; colour = Variety)",
       x = "Light level", y = "Mean quality",
       linetype = "Heat", shape = "Heat", colour = "Variety") +
  theme_minimal(base_size = 13)
#___________________________________________________________________________

anova_tbl <- Anova(fit, type = 2)
anova_tbl

#___________________________________________________________________________

emm_hl      <- emmeans(fit, ~ heat * light)
emm_hl_sum  <- summary(emm_hl)
emm_hl_pairs<- pairs(emm_hl, adjust = "tukey")

kable(emm_hl_sum, digits = 2, caption = "EMMs: Heat × Light")
emm_hl_pairs

#___________________________________________________________________________

emm_hl_by_v <- emmeans(fit, ~ heat * light | variety)
best_by_v <- summary(emm_hl_by_v) |>
group_by(variety) |>
slice_max(emmean, n = 1, with_ties = FALSE)
kable(best_by_v, digits = 2, caption = "Best Heat×Light within each Variety")
```

```r
#___________________________________________________________________________

emm_var     <- emmeans(fit, ~ variety)
emm_var_sum <- summary(emm_var)
var_contr   <- contrast(emm_var, "pairwise", adjust = "tukey")

kable(emm_var_sum, digits = 2, caption = "Variety means (EMMs)")
var_contr

#___________________________________________________________________________


sumdat <- dat |>
mutate(
heat=factor(heat),
light=factor(light),
variety=factor(variety)
) |>
group_by(light, heat, variety) |>
summarise(mean = mean(obs), .groups="drop")

#___________________________________________________________________________

# 1) Mean by Light; lines = Heat; facet Variety

ggplot(sumdat, aes(x=light, y=mean, group=heat, linetype=heat, shape=heat)) +
geom_line() + geom_point(size=2) +
facet_wrap(~ variety) +
labs(title="Mean quality by Light (lines = Heat), faceted by Variety",
x="Light level", y="Mean quality", linetype="Heat", shape="Heat") +
theme_minimal(base_size = 13)

#___________________________________________________________________________

# 2) Mean by Heat; lines = Light; facet Variety

ggplot(sumdat, aes(x=heat, y=mean, group=light,
                   linetype=light, shape=light)) +
geom_line() + geom_point(size=2) +
facet_wrap(~ variety) +
labs(title="Mean quality by Heat (lines = Light), faceted by Variety",
x="Heat level", y="Mean quality", linetype="Light", shape="Light") +
```

```r
theme_minimal(base_size = 13)

#_________________________________________________________________________

# 3) Mean by Light; lines = Heat; colour = Variety

ggplot(sumdat, aes(x=light, y=mean,
group=interaction(heat,variety),
linetype=heat, shape=heat, colour=variety)) +
geom_line() + geom_point(size=2) +
labs(title="Mean quality by Light (lines = Heat; colour = Variety)",
x="Light level", y="Mean quality",
linetype="Heat", shape="Heat", colour="Variety") +
theme_minimal(base_size = 13)

#_________________________________________________________________________

decision_tbl <- emm_hl_sum |>
as.data.frame() |>
mutate(
heat  = as.integer(as.character(heat)),
light = as.integer(as.character(light))
) |>
rename(mean = emmean, lower = lower.CL, upper = upper.CL) |>
arrange(desc(mean)) |>
mutate(best_mean = first(mean),
delta_vs_best = mean - best_mean,   # negative = lower than best
energy_proxy = heat + light) |>
select(heat, light, mean, lower, upper, delta_vs_best, energy_proxy)

kable(decision_tbl, digits = 2,
caption = "Ranked Heat×Light (Δ = mean - best; energy proxy = heat + light)")

near_optimal <- decision_tbl |>
filter(delta_vs_best >= -2.0) |>
arrange(energy_proxy, desc(mean))

kable(near_optimal, digits = 2,
caption = "Near-optimal within 2 points of best (sorted by energy proxy)")

#_________________________________________________________________________
```

```r
# First, create a proper summary with confidence intervals
sumdat_with_ci <- dat %>%
  group_by(light, heat, variety) %>%
  summarise(
    mean = mean(obs),
    n = n(),
    sd = sd(obs),
    se = sd / sqrt(n),
    ci = qt(0.975, df = n-1) * se,  # 95% CI
    .groups = "drop"
  )

# Then find top combinations with proper CIs
top_combos <- sumdat_with_ci %>%
  arrange(desc(mean)) %>%
  mutate(
    mean_quality = mean,
    ci_lower = mean - ci,
    ci_upper = mean + ci
  ) %>%
  head(3)

kable(top_combos[, c("variety", "light", "heat",
                     "mean_quality", "ci_lower", "ci_upper")],
      digits = 2,
      caption = "Top three treatment combinations by mean quality score",
      col.names = c("Variety", "Light", "Heat",
                    "Mean Quality", "CI Lower", "CI Upper"),
      align = c("l", "c", "c", "r", "r", "r"))
```