Experimental Design Project

STA2005S - 2025

Maryam Abrahams (ABRMAR043)
Bheka Mabika (MBKBHE002)

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: Maryam

Figure 1: Blocking example greenhouse

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: There is a significant interaction between chilli variety and the light/heat settings on the quality score. Specifically, the optimal combination of light and heat will differ between the 'Redhot' and 'Furious' varieties.
- H_2: Quality will increase with higher light levels and with intermediate heat levels (2-3), as extreme heat may damage plant cells while insufficient heat limits growth.
- H_3: Both varieties will show similar response patterns to light and heat treatments.
- H_4: 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 2: 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.

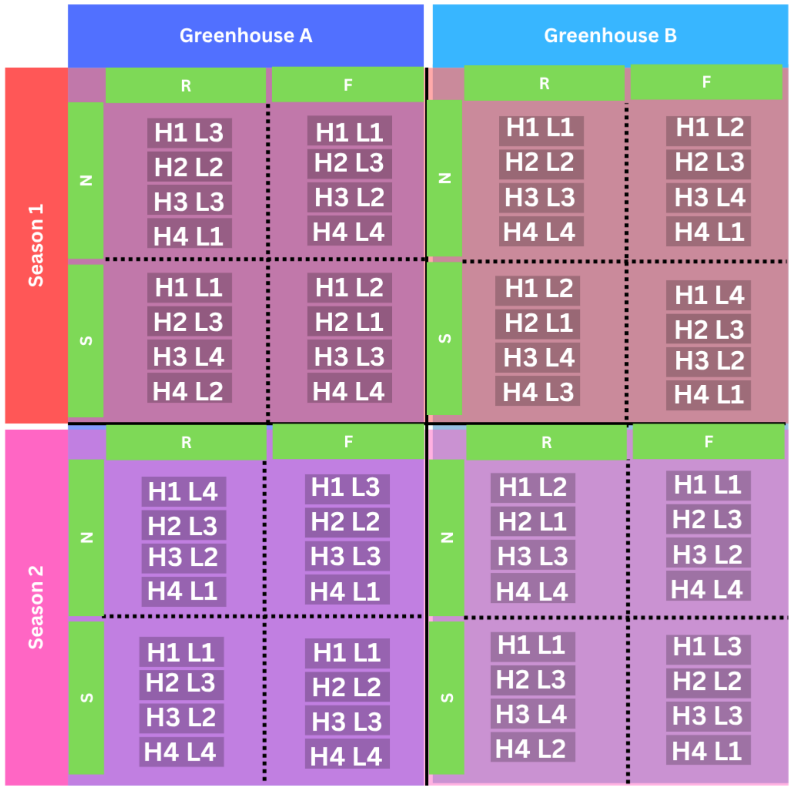

Figure 3: Blocking factors Diagram

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 uses **block** to adjust for background variation across the four season-greenhouse combination positions.

It then tests how **heat**, **light**, and **variety** affect the response and whether they interact.

The **side** factor is included only to correct for small layout or position effects within each greenhouse.

ANOVA:

The experiment used a **Latin Square design** with **Season × Greenhouse** forming the **2×2 blocking structure**, and a full factorial of

Heat (4 levels) × **Light (4 levels)** × **Variety (2 levels)**.

The fitted linear model was:

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

Model fit:

R^2 = 0.968, Adjusted **R^2 = 0.931** → the model explains most of the variation in quality.
Residual SE = 1.76 → small compared to the treatment range.
No clear outliers or leverage points were found when checking diagnostic plots.
Two coefficients were undefined because of overlap between some factors, which is expected in a full factorial design.

Interpretation:

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 ($p = 0.015$) had a small effect, possibly linked to minor position or orientation differences.
All **interaction terms** (e.g., heat $\times$ light) had p-values above 0.1, suggesting the effects act independently.

Overall, the results show that most of the change in quality can be explained by the main treatment factors — especially light — and the blocking effectively reduced environmental variation.

Model Diagnostics:

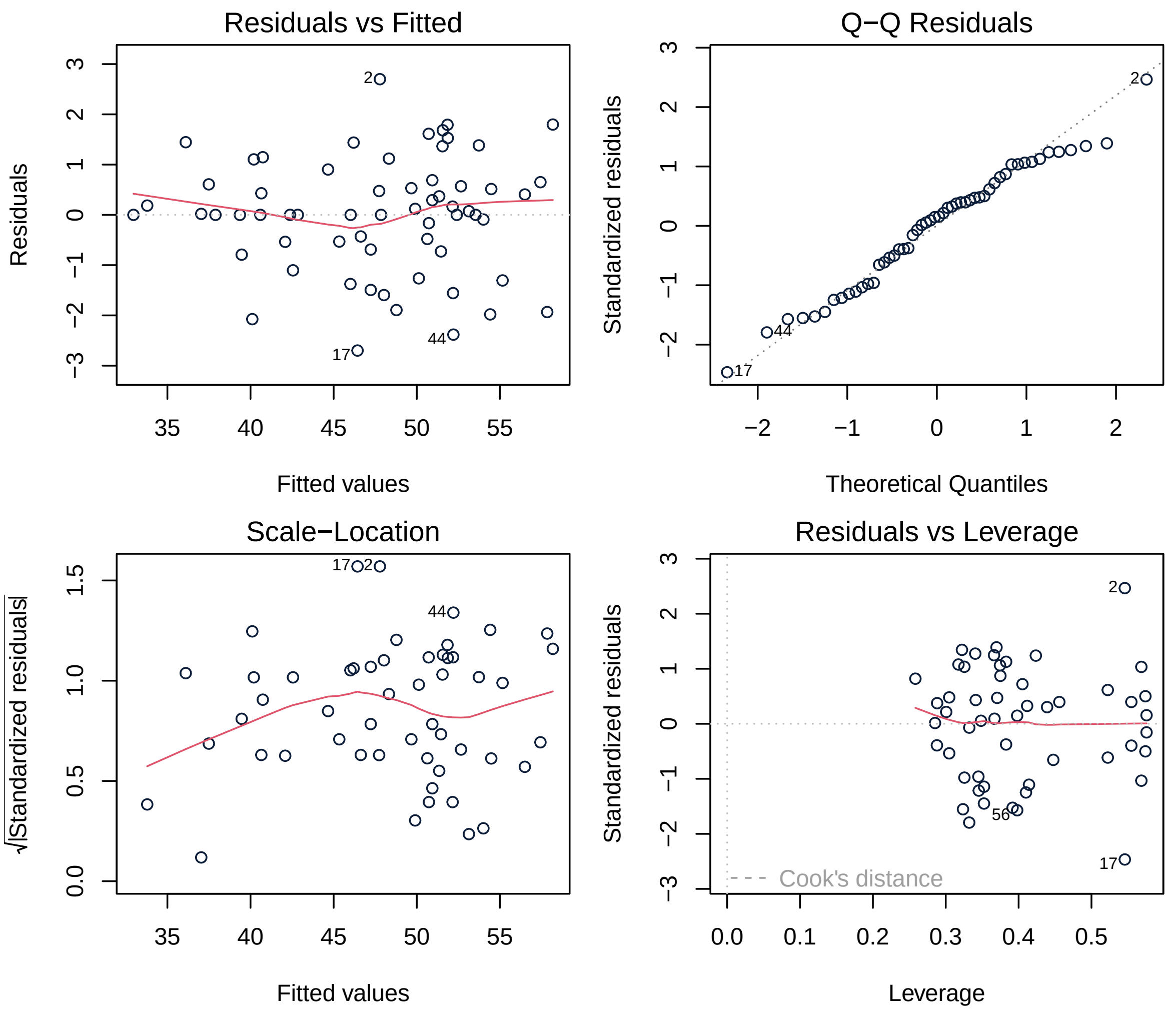

Figure 4: Residual diagnostic plots showing no serious violations of model assumptions. The Q-Q plot indicates approximate normality, and the residuals vs. fitted plot shows no strong patterns suggesting heteroscedasticity.

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:

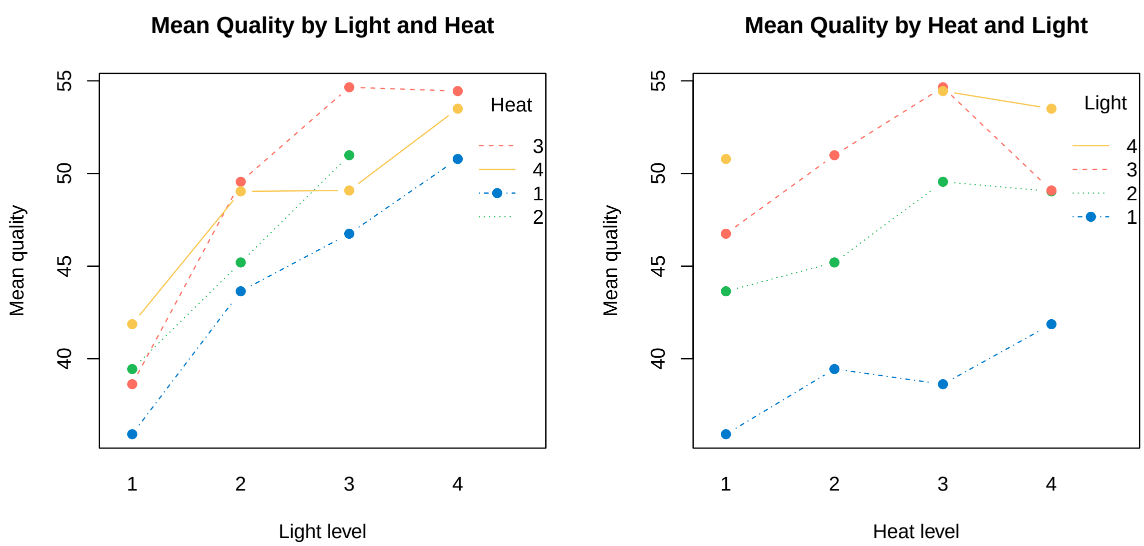

Figure 5: Interaction plots for mean quality based on Heat and Light levels, aggregated across both varieties.

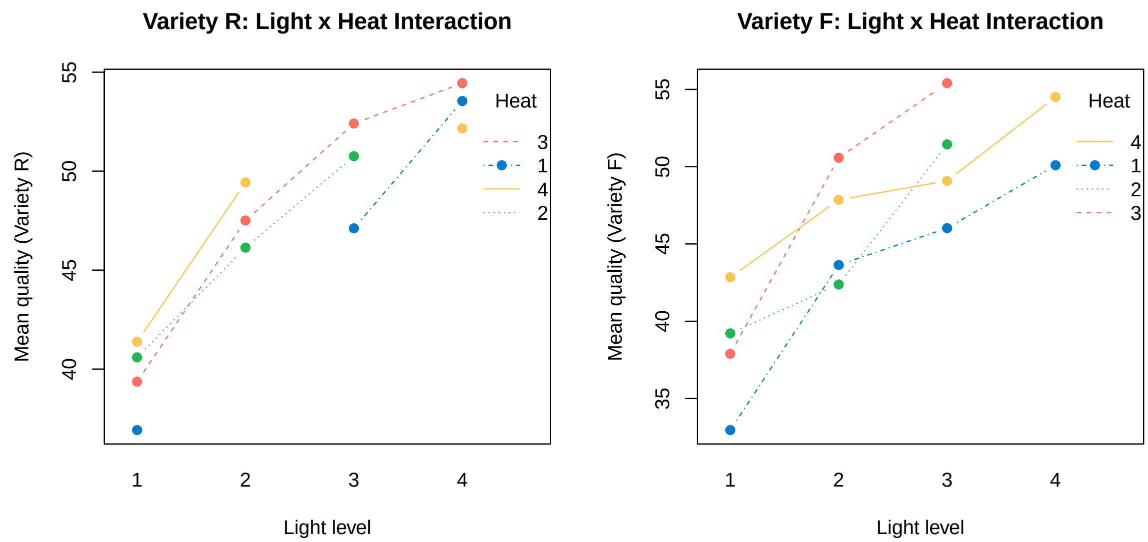

Figure 6: Heat x Light interaction plots for mean quality, shown separately for Variety R (left) and Variety F (right)

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.

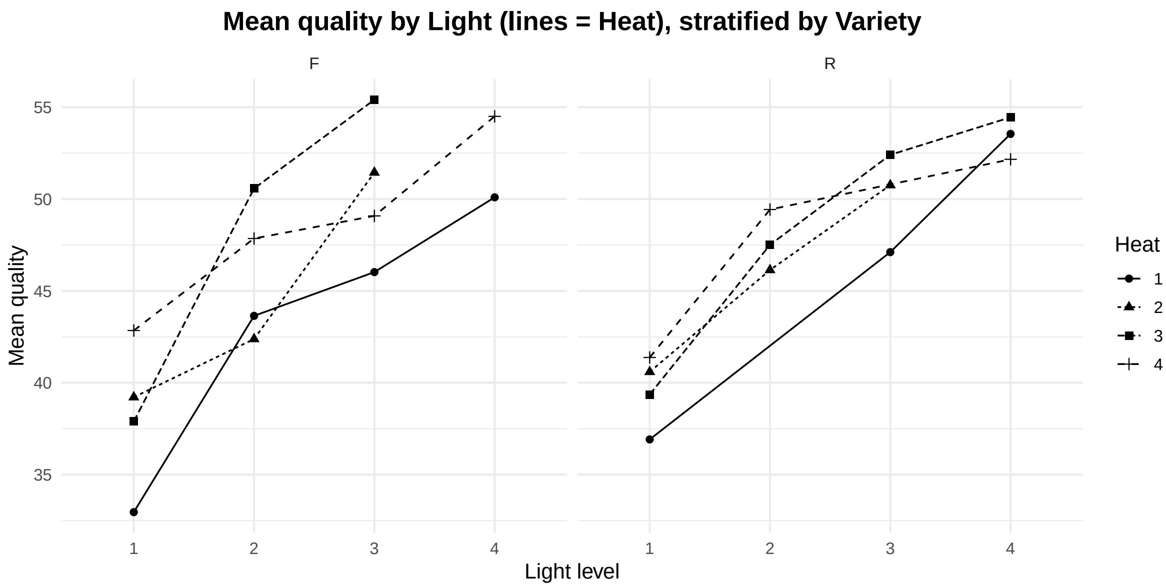

Figure 7: Mean quality stratified by Variety. The plot shows mean quality by Light level (with lines for Heat)

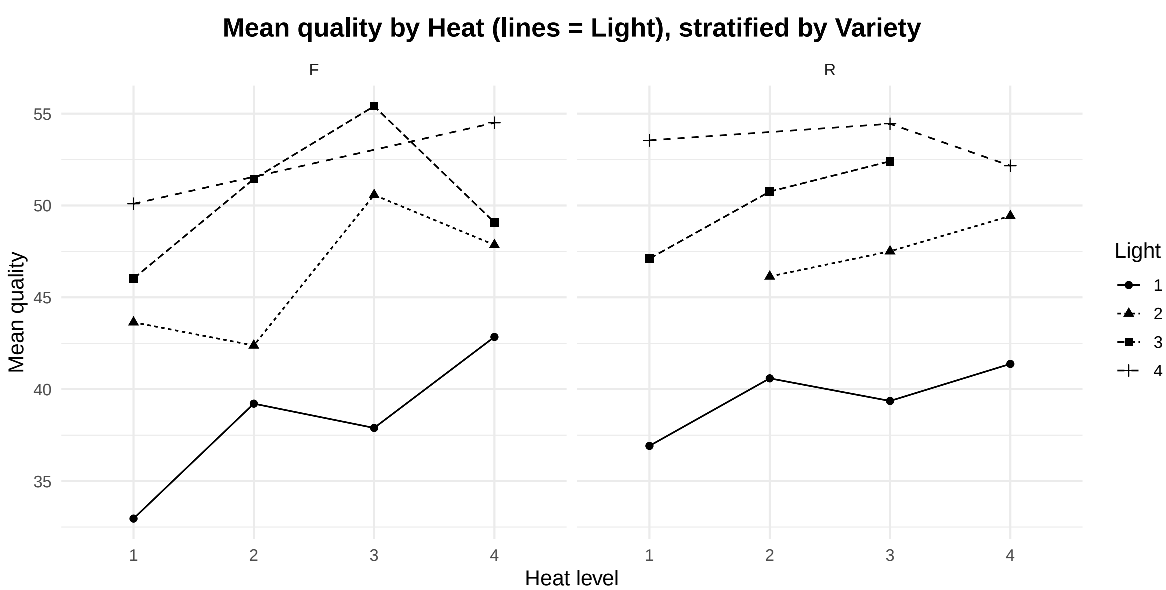

Figure 8: Mean quality stratified by Variety. The plot shows mean quality by Heat level (with lines for Light)

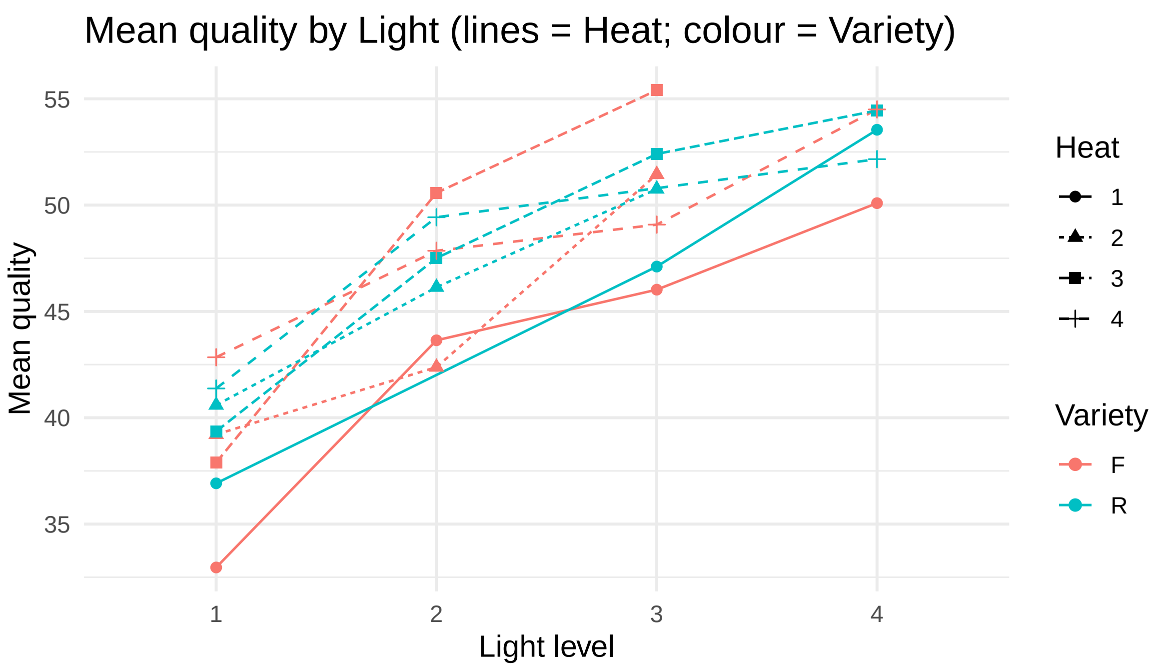

Figure 9: Mean quality by Light level, showing the combined Heat x Light x Variety interaction. Line type represents the Heat level, and color represents the Variety

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                36.05  3   4.5539  0.009335 ** 
heat                215.27  3  27.1898 8.165e-09 ***
light              1456.00  3 183.9003 < 2.2e-16 ***
variety               0.03  1   0.0126  0.911299    
side                153.16  1  58.0354 1.370e-08 ***
heat:light           11.43  8   0.5415  0.816059    
heat:variety         22.09  3   2.7898  0.056931 .  
light:variety         7.31  3   0.9229  0.441268    
heat:light:variety   11.55  5   0.8755  0.508852    
Residuals            81.81 31                       
```

```
---
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	34.81	0.98	31	32.82	36.81
2	1	39.23	0.94	31	37.32	41.13
3	1	40.74	1.22	31	38.26	43.22
4	1	41.26	1.03	31	39.16	43.36
1	2	NA	NA	NA	NA	NA
2	2	45.64	1.00	31	43.59	47.68
3	2	49.17	0.72	31	47.70	50.64
4	2	47.44	0.97	31	45.45	49.42
1	3	48.03	1.08	31	45.83	50.23
2	3	50.97	0.72	31	49.50	52.44
3	3	52.86	0.97	31	50.89	54.83
4	3	NA	NA	NA	NA	NA
1	4	50.65	0.96	31	48.70	52.60
2	4	NA	NA	NA	NA	NA
3	4	NA	NA	NA	NA	NA
4	4	54.04	0.64	31	52.74	55.34

```
contrast                      estimate    SE df t.ratio p.value
heat1 light1 - heat2 light1   -4.412 1.390 31  -3.174  0.1097
heat1 light1 - heat3 light1   -5.927 1.550 31  -3.819  0.0248
heat1 light1 - heat4 light1   -6.446 1.440 31  -4.469  0.0047
heat1 light1 - heat1 light2   nonEst    NA NA      NA      NA
heat1 light1 - heat2 light2  -10.823 1.370 31  -7.896  <.0001
```

```
heat1 light1 - heat3 light2  -14.358 1.220 31 -11.780  <.0001
heat1 light1 - heat4 light2  -12.625 1.410 31  -8.928  <.0001
heat1 light1 - heat1 light3  -13.214 1.450 31  -9.116  <.0001
heat1 light1 - heat2 light3  -16.155 1.210 31 -13.372  <.0001
heat1 light1 - heat3 light3  -18.050 1.380 31 -13.070  <.0001
heat1 light1 - heat4 light3   nonEst    NA NA      NA      NA
heat1 light1 - heat1 light4  -15.836 1.410 31 -11.268  <.0001
heat1 light1 - heat2 light4   nonEst    NA NA      NA      NA
heat1 light1 - heat3 light4   nonEst    NA NA      NA      NA
heat1 light1 - heat4 light4  -19.225 1.150 31 -16.705  <.0001
heat2 light1 - heat3 light1   -1.514 1.590 31  -0.951  0.9978
heat2 light1 - heat4 light1   -2.034 1.340 31  -1.519  0.9239
heat2 light1 - heat1 light2   nonEst    NA NA      NA      NA
heat2 light1 - heat2 light2   -6.410 1.430 31  -4.484  0.0045
heat2 light1 - heat3 light2   -9.946 1.200 31  -8.290  <.0001
heat2 light1 - heat4 light2   -8.212 1.310 31  -6.264  <.0001
heat2 light1 - heat1 light3   -8.801 1.480 31  -5.957  0.0001
heat2 light1 - heat2 light3  -11.743 1.160 31 -10.130  <.0001
heat2 light1 - heat3 light3  -13.638 1.340 31 -10.185  <.0001
heat2 light1 - heat4 light3   nonEst    NA NA      NA      NA
heat2 light1 - heat1 light4  -11.424 1.330 31  -8.607  <.0001
heat2 light1 - heat2 light4   nonEst    NA NA      NA      NA
heat2 light1 - heat3 light4   nonEst    NA NA      NA      NA
heat2 light1 - heat4 light4  -14.813 1.170 31 -12.699  <.0001
heat3 light1 - heat4 light1   -0.520 1.650 31  -0.315  1.0000
heat3 light1 - heat1 light2   nonEst    NA NA      NA      NA
heat3 light1 - heat2 light2   -4.896 1.500 31  -3.266  0.0899
heat3 light1 - heat3 light2   -8.432 1.380 31  -6.115  0.0001
heat3 light1 - heat4 light2   -6.698 1.590 31  -4.207  0.0093
heat3 light1 - heat1 light3   -7.287 1.560 31  -4.677  0.0027
heat3 light1 - heat2 light3  -10.228 1.450 31  -7.072  <.0001
heat3 light1 - heat3 light3  -12.123 1.560 31  -7.784  <.0001
heat3 light1 - heat4 light3   nonEst    NA NA      NA      NA
heat3 light1 - heat1 light4   -9.909 1.540 31  -6.444  <.0001
heat3 light1 - heat2 light4   nonEst    NA NA      NA      NA
heat3 light1 - heat3 light4   nonEst    NA NA      NA      NA
heat3 light1 - heat4 light4  -13.299 1.330 31  -9.978  <.0001
heat4 light1 - heat1 light2   nonEst    NA NA      NA      NA
heat4 light1 - heat2 light2   -4.376 1.500 31  -2.926  0.1807
heat4 light1 - heat3 light2   -7.912 1.280 31  -6.200  <.0001
heat4 light1 - heat4 light2   -6.178 1.380 31  -4.484  0.0045
heat4 light1 - heat1 light3   -6.767 1.520 31  -4.458  0.0048
heat4 light1 - heat2 light3   -9.709 1.230 31  -7.881  <.0001
```

```
heat4 light1 - heat3 light3  -11.604 1.420 31  -8.190  <.0001
heat4 light1 - heat4 light3   nonEst    NA NA      NA      NA
heat4 light1 - heat1 light4   -9.390 1.400 31  -6.712  <.0001
heat4 light1 - heat2 light4   nonEst    NA NA      NA      NA
heat4 light1 - heat3 light4   nonEst    NA NA      NA      NA
heat4 light1 - heat4 light4  -12.779 1.240 31 -10.320  <.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   -3.536 1.200 31  -2.939  0.1763
heat2 light2 - heat4 light2   -1.802 1.450 31  -1.243  0.9805
heat2 light2 - heat1 light3   -2.391 1.450 31  -1.647  0.8778
heat2 light2 - heat2 light3   -5.332 1.260 31  -4.221  0.0090
heat2 light2 - heat3 light3   -7.227 1.380 31  -5.250  0.0006
heat2 light2 - heat4 light3   nonEst    NA NA      NA      NA
heat2 light2 - heat1 light4   -5.013 1.390 31  -3.598  0.0423
heat2 light2 - heat2 light4   nonEst    NA NA      NA      NA
heat2 light2 - heat3 light4   nonEst    NA NA      NA      NA
heat2 light2 - heat4 light4   -8.403 1.150 31  -7.310  <.0001
heat3 light2 - heat4 light2    1.734 1.220 31   1.426  0.9489
heat3 light2 - heat1 light3    1.145 1.280 31   0.894  0.9987
heat3 light2 - heat2 light3   -1.797 1.040 31  -1.729  0.8413
heat3 light2 - heat3 light3   -3.692 1.190 31  -3.099  0.1281
heat3 light2 - heat4 light3   nonEst    NA NA      NA      NA
heat3 light2 - heat1 light4   -1.478 1.170 31  -1.260  0.9784
heat3 light2 - heat2 light4   nonEst    NA NA      NA      NA
heat3 light2 - heat3 light4   nonEst    NA NA      NA      NA
heat3 light2 - heat4 light4   -4.867 0.947 31  -5.140  0.0008
heat4 light2 - heat1 light3   -0.589 1.450 31  -0.405  1.0000
heat4 light2 - heat2 light3   -3.530 1.200 31  -2.933  0.1781
heat4 light2 - heat3 light3   -5.425 1.370 31  -3.957  0.0176
heat4 light2 - heat4 light3   nonEst    NA NA      NA      NA
heat4 light2 - heat1 light4   -3.211 1.330 31  -2.420  0.4229
heat4 light2 - heat2 light4   nonEst    NA NA      NA      NA
heat4 light2 - heat3 light4   nonEst    NA NA      NA      NA
```

```
 heat4 light2 - heat4 light4   -6.601 1.190 31  -5.568  0.0002
 heat1 light3 - heat2 light3   -2.941 1.310 31  -2.241  0.5350
 heat1 light3 - heat3 light3   -4.836 1.490 31  -3.242  0.0949
 heat1 light3 - heat4 light3   nonEst    NA NA      NA      NA
 heat1 light3 - heat1 light4   -2.622 1.430 31  -1.830  0.7900
 heat1 light3 - heat2 light4   nonEst    NA NA      NA      NA
 heat1 light3 - heat3 light4   nonEst    NA NA      NA      NA
 heat1 light3 - heat4 light4   -6.012 1.220 31  -4.925  0.0014
 heat2 light3 - heat3 light3   -1.895 1.220 31  -1.556  0.9119
 heat2 light3 - heat4 light3   nonEst    NA NA      NA      NA
 heat2 light3 - heat1 light4    0.319 1.220 31   0.262  1.0000
 heat2 light3 - heat2 light4   nonEst    NA NA      NA      NA
 heat2 light3 - heat3 light4   nonEst    NA NA      NA      NA
 heat2 light3 - heat4 light4   -3.070 0.975 31  -3.148  0.1156
 heat3 light3 - heat4 light3   nonEst    NA NA      NA      NA
 heat3 light3 - heat1 light4    2.214 1.330 31   1.659  0.8727
 heat3 light3 - heat2 light4   nonEst    NA NA      NA      NA
 heat3 light3 - heat3 light4   nonEst    NA NA      NA      NA
 heat3 light3 - heat4 light4   -1.175 1.160 31  -1.010  0.9962
 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   -3.389 1.150 31  -2.936  0.1772
 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: block, variety, side
P value adjustment: tukey method for comparing a family of 12 estimates
```

Answer: Highest mean quality is at **Heat 3 $\times$ Light 4**. Next best are typically **H4$\times$L3** and **H3$\times$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	54.93	0.85	31	53.21	56.66
3	4	R	54.19	0.87	31	52.42	55.95

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: block, heat, light, side
```

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

Mean plots (no CIs; simple patterns)

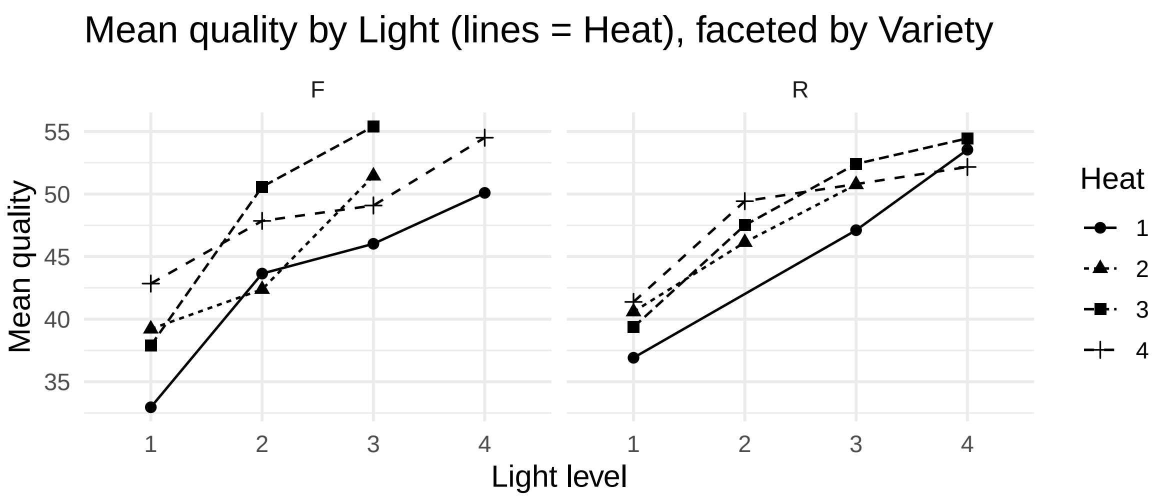

Figure 10: Mean plot with Light × Heat faceted by variety. Light shows strong positive effect on quality across all heat levels and both varieties.

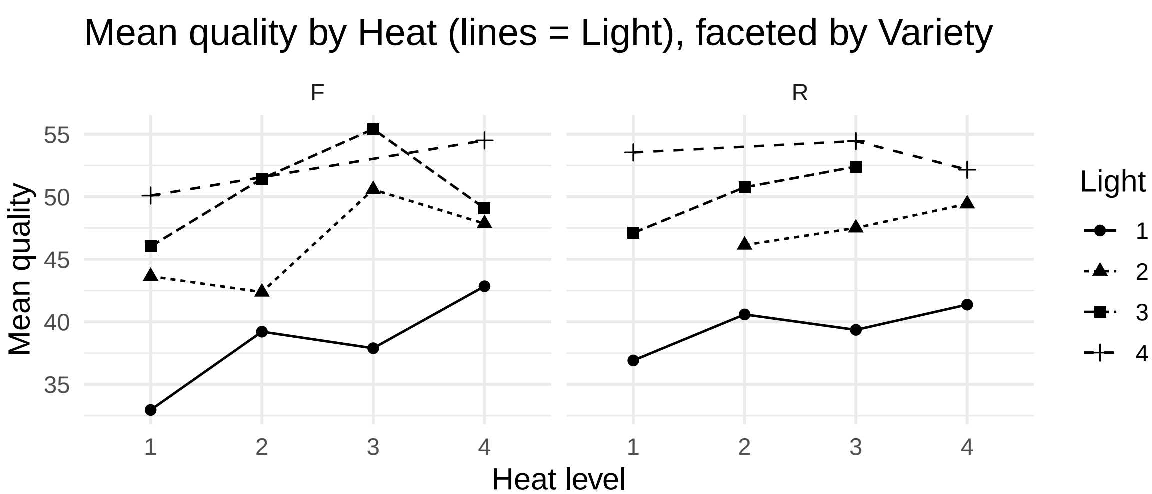

Figure 11: Mean plot with Heat × Light faceted by variety.Heat effects are weaker and less consistent than light effects, with similar patterns across varieties.

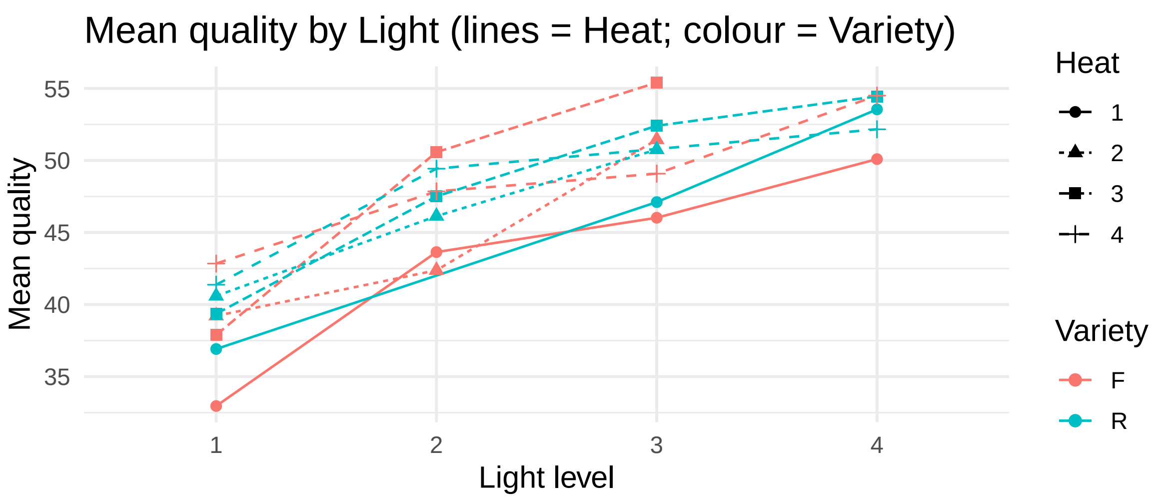

Figure 12: Three-way imean plot of Light × Heat × Variety. Parallel response patterns suggest minimal interactions; light dominates quality improvements.

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	54.04	52.74	55.34	0.00	8
3	3	52.86	50.89	54.83	-1.18	6
2	3	50.97	49.50	52.44	-3.07	5
1	4	50.65	48.70	52.60	-3.39	5
3	2	49.17	47.70	50.64	-4.87	5
1	3	48.03	45.83	50.23	-6.01	4
4	2	47.44	45.45	49.42	-6.60	6
2	2	45.64	43.59	47.68	-8.40	4
4	1	41.26	39.16	43.36	-12.78	5
3	1	40.74	38.26	43.22	-13.30	4
2	1	39.23	37.32	41.13	-14.81	3

heat	light	mean	lower	upper	delta_vs_best	energy_proxy
1	1	34.81	32.82	36.81	-19.23	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	52.86	50.89	54.83	-1.18	6
4	4	54.04	52.74	55.34	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)
