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.

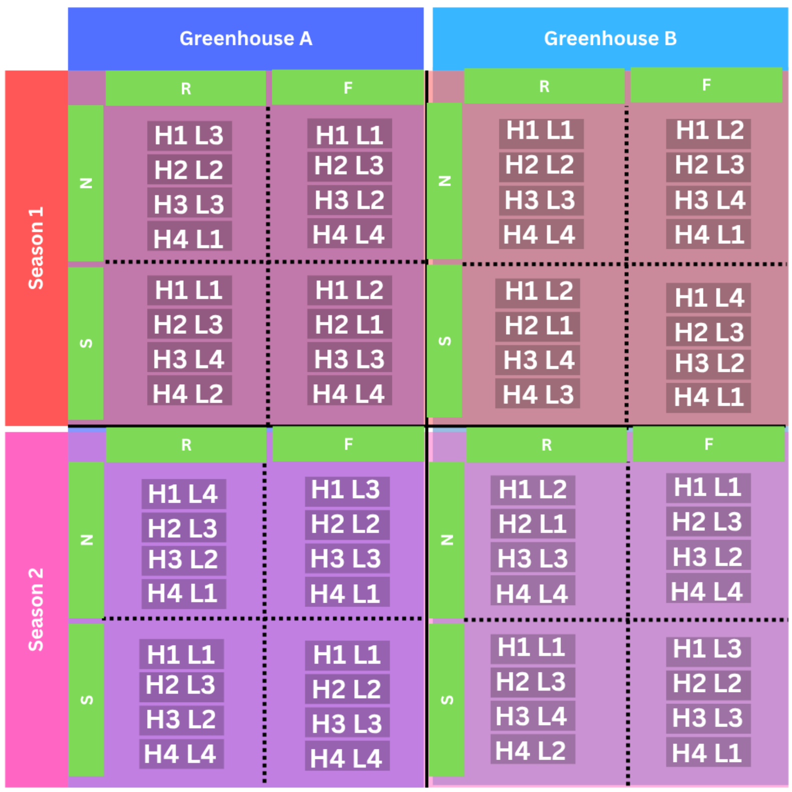

Figure 2: Blocking factors Diagram

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 experimental design, using **Season** × **Greenhouse** (`block`) to control for background variation. The model tested the main effects of **heat**, **light**, **variety**, and **side**, as well as all two-way interactions involving the primary treatment factors (`heat:light`, `heat:variety`, `light:variety`).

Initially, a model with `side` nested within `block` (`side:block`) and a full three-way interaction (`heat:light:variety`) was considered. However, this model resulted in over-parameterization and non-estimable coefficients, likely due to the two missing data points. Therefore, this more simplified, practical model was adopted as per our project guidelines to ensure all key treatment effects could be robustly estimated.

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.

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** (`block`) as blocks. To ensure model stability, a simplified fixed-effects model was fitted. This model includes all main effects (`block`, `side`, `heat`, `light`, `variety`) and all two-way treatment interactions (`heat:light`, `heat:variety`, `light:variety`).

The fitted linear model was: `obs ~ block + side + heat*light + variety + heat:variety + light:variety`

$$\begin{aligned}\text{obs} =& \text{block} + \text{side} + \text{heat} + \text{light} + \text{variety}\\ &+ (\text{heat} : \text{light}) + (\text{heat} : \text{variety}) + (\text{light} : \text{variety})\end{aligned}$$

Table 2: Type II ANOVA Results

	Sum Sq	Df	F value	Pr(>F)
block	43.9712658	3	5.6515190	0.0028015
side	162.9796258	1	62.8421158	0.0000000
heat	215.2706776	3	27.6682127	0.0000000
light	1455.9992346	3	187.1360137	0.0000000
variety	0.0332905	1	0.0128362	0.9104245
heat:light	11.4327538	8	0.5510339	0.8099978
heat:variety	22.0876775	3	2.8388751	0.0515297
light:variety	7.3069467	3	0.9391440	0.4318596
Residuals	93.3651972	36	NA	NA

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.

Model Diagnostics:

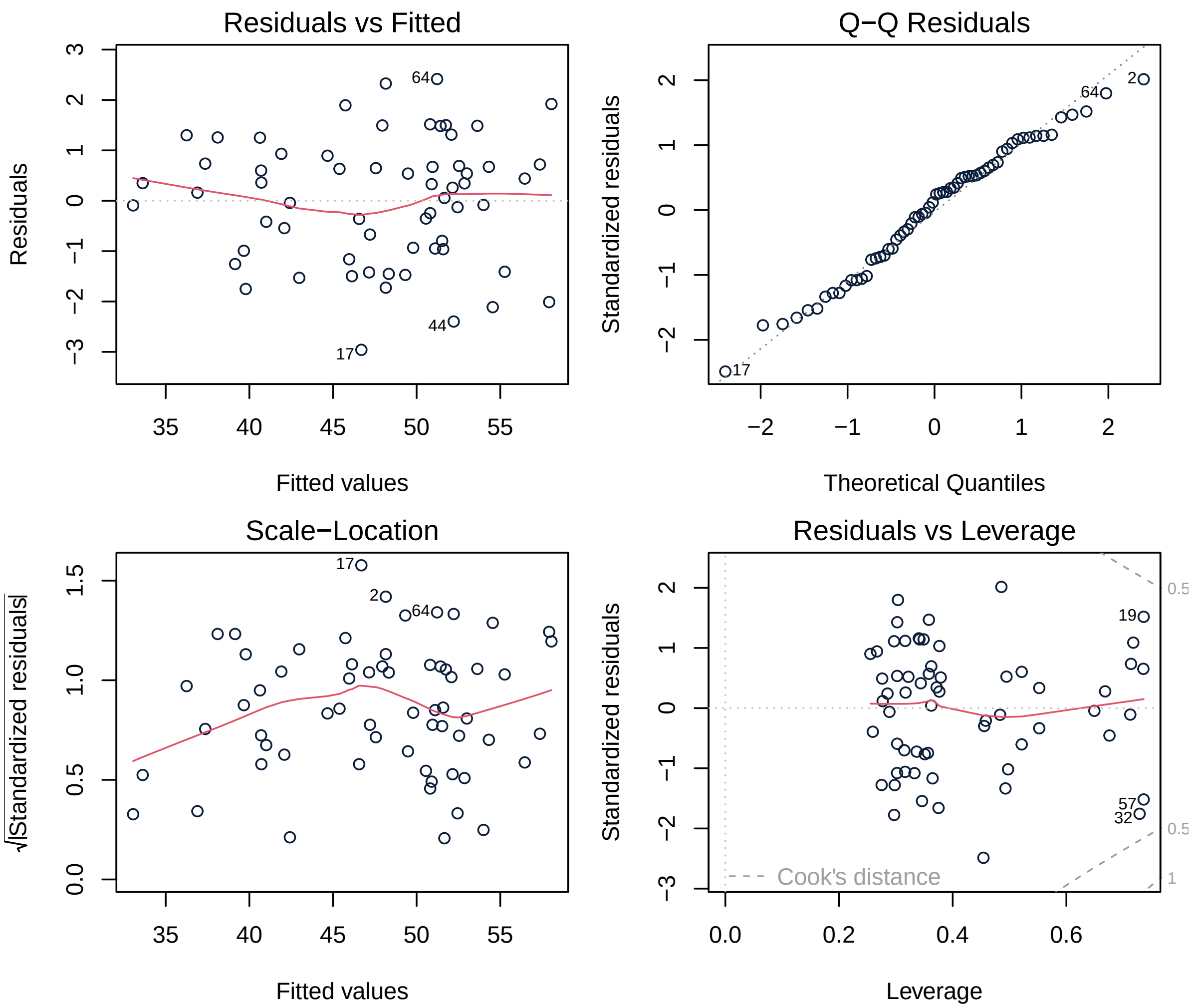

Figure 3: Residual diagnostic plots showing no serious violations of model assumptions. The Q-Q plot indicates approximate normality, and the residuals vs. fitted plot shows no strong patterns suggesting heteroscedasticity.

The model assumptions were checked using standard residual plots. The Residuals vs Fitted plot shows a random scatter of points around zero, indicating that the assumption of linearity and homoscedasticity (constant variance) is satisfied. There are no visible patterns or funnel shapes.

The Q-Q plot shows that the residuals lie close to the straight line, confirming that the normality assumption is reasonable. The Scale-Location plot also shows roughly constant spread across fitted values, supporting equal variance.

Finally, the residuals vs leverage plot shows that no points have unusually high leverage or cooks distance, indicating that no single observation unduly influenced the fitted model.

Overall, these diagnostic plots confirm that the model fits the data well, and the results from the type 2 anova can be considered reliable and valid for interpretation.

Interaction 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.

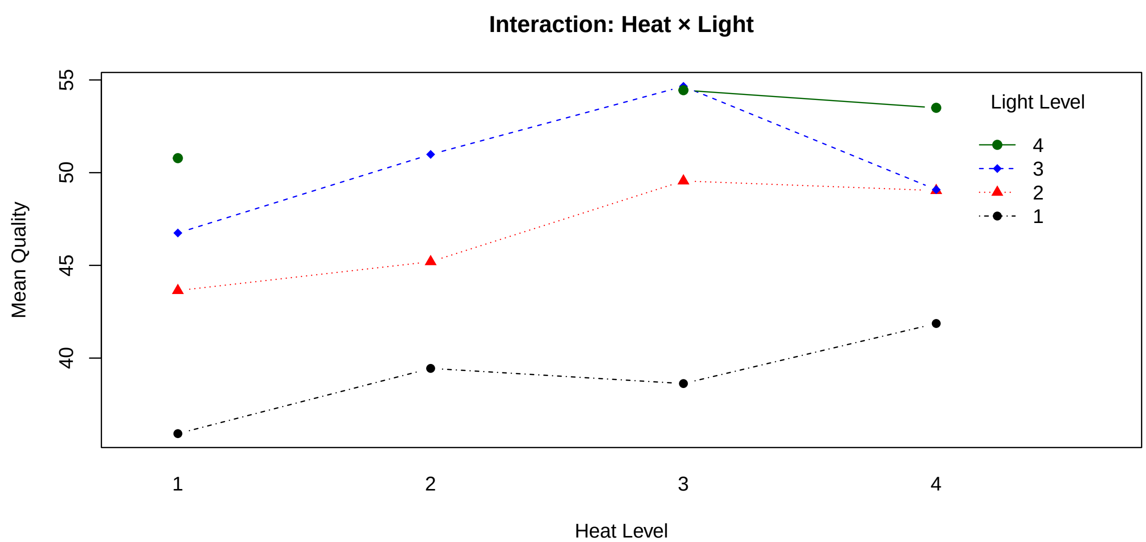

Figure 4: Interaction between Heat and Light levels on chilli quality. Lines represent different Light levels.

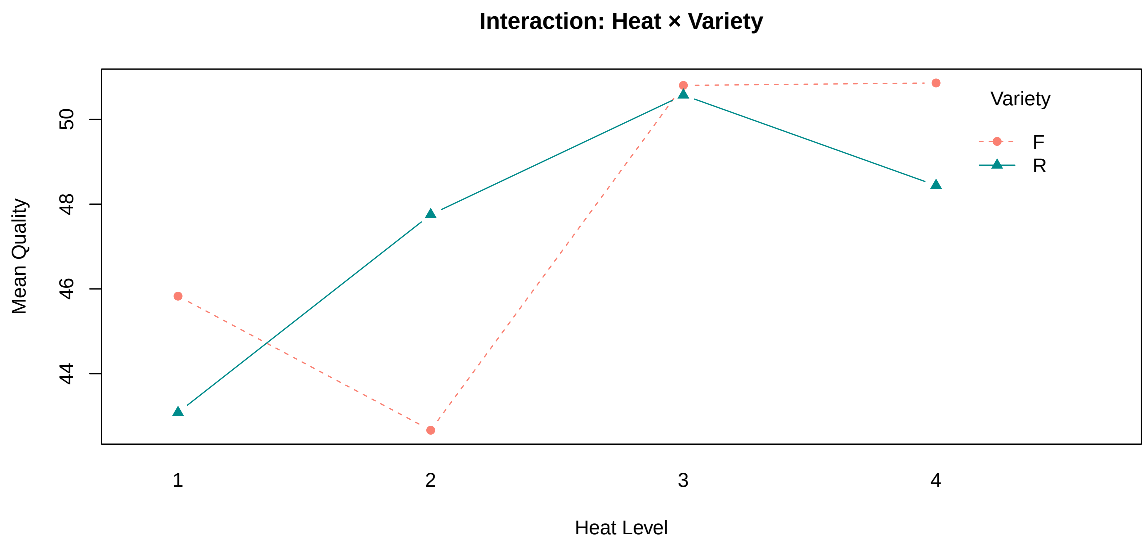

Figure 5: Interaction between Heat levels and Variety on chilli quality. Lines show how Redhot and Furious varieties respond differently to temperature.

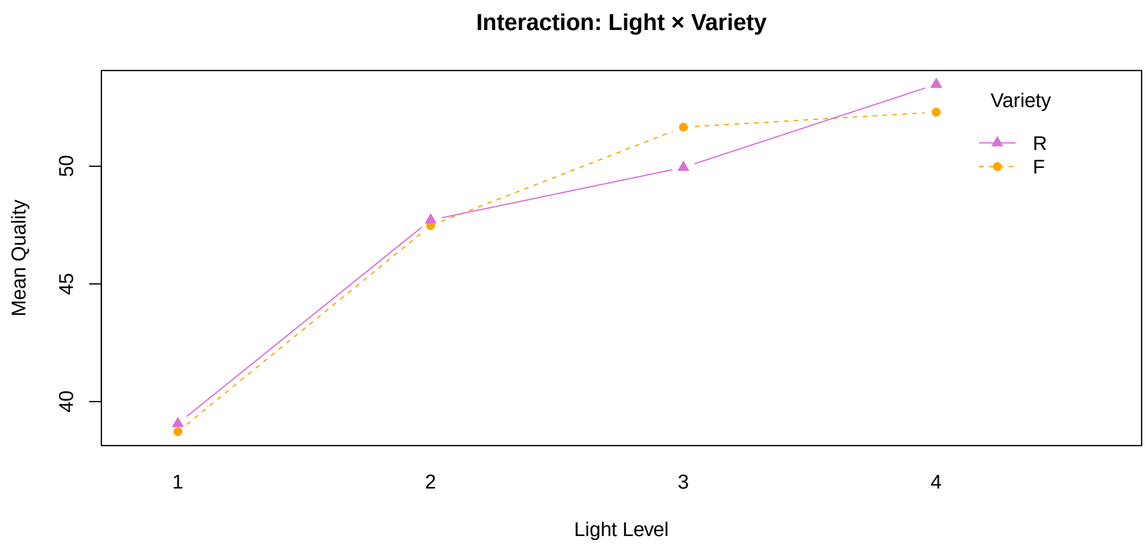

Figure 6: Interaction between Heat levels and Variety on chilli quality. Lines show how Redhot and Furious varieties respond differently to temperature.

Interpretation of Interaction Plots

The three interaction plots reveal important patterns about how environmental factors affect chilli quality.

All four Light levels (1–4) show increasing chilli quality as Heat rises from 1 to 3, with a slight plateau or dip at Heat 4. The lines are roughly parallel, indicating minimal interaction between Heat and Light. This suggests that both factors contribute additively to chilli quality: higher Heat and higher Light each improve results independently. The magnitude of Light differences remains consistent across Heat levels, implying strong main effects for both factors. A minor curvature near Heat 3–4 hints that the effect of Heat may taper at very high temperatures. Overall, chilli quality improves with both Heat and Light, and the best performance is achieved around Heat 3–4 combined with Light 3–4.

Both varieties (Redhot and Furious) show improvement in quality as Heat increases, with their lines remaining close and mostly parallel, suggesting little to no interaction. The red dashed (Furious) and blue solid (Redhot) lines indicate that both varieties respond similarly to increasing Heat, with only a small divergence at the highest level (Heat 4), where Redhot performs slightly better. This pattern implies that the Heat effect is consistent across varieties, confirming no meaningful Heat × Variety interaction. Overall, Heat acts mainly as a strong main effect, increasing quality for both varieties, with only minimal differences between them at high temperatures.

Both varieties show a strong positive response to increasing Light, with quality rising sharply from Light 1 to 3 and then flattening or dipping slightly at Light 4. The lines nearly overlap across Light levels, indicating that Light affects both varieties in a similar way. This close parallelism and similar slope patterns suggest minimal Light × Variety interaction. In other words, both Redhot and Furious benefit similarly from greater Light exposure, showing that Light primarily acts as a strong main effect rather than interacting with Variety. Overall, both varieties achieve top quality under Light levels 3–4.

Conclusions:

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

```
 heat light emmean    SE df lower.CL upper.CL
 1    1       34.9 0.913 36     33.0     36.7
 2    1       39.3 0.819 36     37.6     40.9
```

```
 3    1       40.9 1.200 36     38.5     43.3
 4    1       40.9 0.981 36     38.9     42.9
 1    2       44.9 1.460 36     41.9     47.8
 2    2       45.8 0.927 36     44.0     47.7
 3    2       49.4 0.694 36     48.0     50.8
 4    2       47.8 0.926 36     46.0     49.7
 1    3       47.9 1.030 36     45.8     50.0
 2    3       50.9 0.693 36     49.5     52.3
 3    3       52.9 0.889 36     51.1     54.7
 4    3       49.3 1.580 36     46.1     52.6
 1    4       50.5 0.898 36     48.7     52.3
 2    4     nonEst    NA NA       NA       NA
 3    4       55.8 1.330 36     53.1     58.5
 4    4       54.1 0.629 36     52.8     55.4

Results are averaged over the levels of: block, side, variety
Confidence level used: 0.95

 contrast                    estimate    SE df t.ratio p.value
 heat1 light1 - heat2 light1  -4.3723 1.310 36  -3.348  0.0975
 heat1 light1 - heat3 light1  -6.0192 1.480 36  -4.053  0.0173
 heat1 light1 - heat4 light1  -5.9985 1.350 36  -4.449  0.0059
 heat1 light1 - heat1 light2  -9.9625 1.820 36  -5.476  0.0003
 heat1 light1 - heat2 light2 -10.9409 1.250 36  -8.719  <.0001
 heat1 light1 - heat3 light2 -14.5451 1.140 36 -12.728  <.0001
 heat1 light1 - heat4 light2 -12.9547 1.340 36  -9.643  <.0001
 heat1 light1 - heat1 light3 -13.0108 1.360 36  -9.569  <.0001
 heat1 light1 - heat2 light3 -16.0448 1.160 36 -13.783  <.0001
 heat1 light1 - heat3 light3 -17.9597 1.250 36 -14.352  <.0001
 heat1 light1 - heat4 light3 -14.4567 1.730 36  -8.363  <.0001
 heat1 light1 - heat1 light4 -15.6040 1.350 36 -11.601  <.0001
 heat1 light1 - heat2 light4   nonEst    NA NA      NA      NA
 heat1 light1 - heat3 light4 -20.8996 1.650 36 -12.688  <.0001
 heat1 light1 - heat4 light4 -19.2295 1.090 36 -17.717  <.0001
 heat2 light1 - heat3 light1  -1.6469 1.480 36  -1.113  0.9979
 heat2 light1 - heat4 light1  -1.6262 1.270 36  -1.281  0.9919
 heat2 light1 - heat1 light2  -5.5902 1.710 36  -3.275  0.1146
 heat2 light1 - heat2 light2  -6.5687 1.330 36  -4.946  0.0014
 heat2 light1 - heat3 light2 -10.1729 1.080 36  -9.410  <.0001
 heat2 light1 - heat4 light2  -8.5825 1.150 36  -7.446  <.0001
 heat2 light1 - heat1 light3  -8.6386 1.300 36  -6.663  <.0001
 heat2 light1 - heat2 light3 -11.6726 1.100 36 -10.658  <.0001
```

```
heat2 light1 - heat3 light3 -13.5875 1.230 36 -11.075  <.0001
heat2 light1 - heat4 light3 -10.0844 1.910 36  -5.278  0.0005
heat2 light1 - heat1 light4 -11.2318 1.180 36  -9.531  <.0001
heat2 light1 - heat2 light4   nonEst    NA NA      NA      NA
heat2 light1 - heat3 light4 -16.5274 1.610 36 -10.261  <.0001
heat2 light1 - heat4 light4 -14.8572 1.060 36 -14.070  <.0001
heat3 light1 - heat4 light1   0.0207 1.580 36   0.013  1.0000
heat3 light1 - heat1 light2  -3.9433 1.950 36  -2.021  0.7735
heat3 light1 - heat2 light2  -4.9217 1.450 36  -3.403  0.0863
heat3 light1 - heat3 light2  -8.5260 1.360 36  -6.254  <.0001
heat3 light1 - heat4 light2  -6.9355 1.560 36  -4.438  0.0061
heat3 light1 - heat1 light3  -6.9917 1.530 36  -4.578  0.0041
heat3 light1 - heat2 light3 -10.0257 1.400 36  -7.151  <.0001
heat3 light1 - heat3 light3 -11.9405 1.490 36  -8.015  <.0001
heat3 light1 - heat4 light3  -8.4375 1.910 36  -4.416  0.0065
heat3 light1 - heat1 light4  -9.5848 1.490 36  -6.421  <.0001
heat3 light1 - heat2 light4   nonEst    NA NA      NA      NA
heat3 light1 - heat3 light4 -14.8804 1.860 36  -7.992  <.0001
heat3 light1 - heat4 light4 -13.2103 1.320 36 -10.030  <.0001
heat4 light1 - heat1 light2  -3.9640 1.720 36  -2.308  0.5916
heat4 light1 - heat2 light2  -4.9425 1.410 36  -3.512  0.0671
heat4 light1 - heat3 light2  -8.5467 1.200 36  -7.104  <.0001
heat4 light1 - heat4 light2  -6.9563 1.280 36  -5.449  0.0003
heat4 light1 - heat1 light3  -7.0124 1.420 36  -4.943  0.0014
heat4 light1 - heat2 light3 -10.0464 1.190 36  -8.431  <.0001
heat4 light1 - heat3 light3 -11.9612 1.330 36  -8.988  <.0001
heat4 light1 - heat4 light3  -8.4582 1.990 36  -4.241  0.0104
heat4 light1 - heat1 light4  -9.6055 1.290 36  -7.426  <.0001
heat4 light1 - heat2 light4   nonEst    NA NA      NA      NA
heat4 light1 - heat3 light4 -14.9012 1.670 36  -8.900  <.0001
heat4 light1 - heat4 light4 -13.2310 1.190 36 -11.144  <.0001
heat1 light2 - heat2 light2  -0.9785 1.850 36  -0.528  1.0000
heat1 light2 - heat3 light2  -4.5827 1.610 36  -2.852  0.2661
heat1 light2 - heat4 light2  -2.9923 1.790 36  -1.676  0.9271
heat1 light2 - heat1 light3  -3.0484 1.870 36  -1.631  0.9399
heat1 light2 - heat2 light3  -6.0824 1.570 36  -3.864  0.0283
heat1 light2 - heat3 light3  -7.9972 1.780 36  -4.501  0.0051
heat1 light2 - heat4 light3  -4.4942 2.240 36  -2.005  0.7823
heat1 light2 - heat1 light4  -5.6416 1.670 36  -3.376  0.0916
heat1 light2 - heat2 light4   nonEst    NA NA      NA      NA
heat1 light2 - heat3 light4 -10.9372 1.720 36  -6.366  <.0001
heat1 light2 - heat4 light4  -9.2670 1.620 36  -5.718  0.0001
heat2 light2 - heat3 light2  -3.6042 1.160 36  -3.103  0.1646
```

```
heat2 light2 - heat4 light2  -2.0138 1.340 36  -1.498  0.9685
heat2 light2 - heat1 light3  -2.0699 1.370 36  -1.511  0.9664
heat2 light2 - heat2 light3  -5.1039 1.160 36  -4.401  0.0067
heat2 light2 - heat3 light3  -7.0188 1.240 36  -5.657  0.0002
heat2 light2 - heat4 light3  -3.5157 1.680 36  -2.093  0.7309
heat2 light2 - heat1 light4  -4.6631 1.310 36  -3.559  0.0600
heat2 light2 - heat2 light4   nonEst    NA NA      NA      NA
heat2 light2 - heat3 light4  -9.9587 1.690 36  -5.897  0.0001
heat2 light2 - heat4 light4  -8.2886 1.080 36  -7.650  <.0001
heat3 light2 - heat4 light2   1.5904 1.180 36   1.347  0.9872
heat3 light2 - heat1 light3   1.5343 1.230 36   1.247  0.9937
heat3 light2 - heat2 light3  -1.4997 0.985 36  -1.523  0.9641
heat3 light2 - heat3 light3  -3.4146 1.100 36  -3.113  0.1614
heat3 light2 - heat4 light3   0.0885 1.730 36   0.051  1.0000
heat3 light2 - heat1 light4  -1.0589 1.110 36  -0.957  0.9996
heat3 light2 - heat2 light4   nonEst    NA NA      NA      NA
heat3 light2 - heat3 light4  -6.3545 1.590 36  -4.003  0.0197
heat3 light2 - heat4 light4  -4.6843 0.930 36  -5.035  0.0011
heat4 light2 - heat1 light3  -0.0561 1.390 36  -0.041  1.0000
heat4 light2 - heat2 light3  -3.0901 1.150 36  -2.691  0.3492
heat4 light2 - heat3 light3  -5.0050 1.270 36  -3.942  0.0231
heat4 light2 - heat4 light3  -1.5019 2.010 36  -0.747  1.0000
heat4 light2 - heat1 light4  -2.6493 1.230 36  -2.151  0.6943
heat4 light2 - heat2 light4   nonEst    NA NA      NA      NA
heat4 light2 - heat3 light4  -7.9449 1.710 36  -4.646  0.0034
heat4 light2 - heat4 light4  -6.2748 1.150 36  -5.438  0.0003
heat1 light3 - heat2 light3  -3.0340 1.230 36  -2.466  0.4871
heat1 light3 - heat3 light3  -4.9489 1.400 36  -3.531  0.0642
heat1 light3 - heat4 light3  -1.4458 1.990 36  -0.728  1.0000
heat1 light3 - heat1 light4  -2.5932 1.410 36  -1.843  0.8648
heat1 light3 - heat2 light4   nonEst    NA NA      NA      NA
heat1 light3 - heat3 light4  -7.8888 1.750 36  -4.498  0.0051
heat1 light3 - heat4 light4  -6.2186 1.180 36  -5.288  0.0005
heat2 light3 - heat3 light3  -1.9149 1.160 36  -1.652  0.9342
heat2 light3 - heat4 light3   1.5882 1.790 36   0.886  0.9998
heat2 light3 - heat1 light4   0.4408 1.150 36   0.384  1.0000
heat2 light3 - heat2 light4   nonEst    NA NA      NA      NA
heat2 light3 - heat3 light4  -4.8548 1.470 36  -3.305  0.1074
heat2 light3 - heat4 light4  -3.1846 0.944 36  -3.372  0.0924
heat3 light3 - heat4 light3   3.5031 1.720 36   2.042  0.7613
heat3 light3 - heat1 light4   2.3557 1.240 36   1.893  0.8416
heat3 light3 - heat2 light4   nonEst    NA NA      NA      NA
heat3 light3 - heat3 light4  -2.9399 1.710 36  -1.717  0.9138
```

```
 heat3 light3 - heat4 light4  -1.2698 1.090 36  -1.164  0.9968
 heat4 light3 - heat1 light4  -1.1474 1.890 36  -0.606  1.0000
 heat4 light3 - heat2 light4   nonEst    NA NA      NA      NA
 heat4 light3 - heat3 light4  -6.4430 1.980 36  -3.254  0.1199
 heat4 light3 - heat4 light4  -4.7728 1.660 36  -2.870  0.2575
 heat1 light4 - heat2 light4   nonEst    NA NA      NA      NA
 heat1 light4 - heat3 light4  -5.2956 1.800 36  -2.935  0.2284
 heat1 light4 - heat4 light4  -3.6255 1.100 36  -3.303  0.1079
 heat2 light4 - heat3 light4   nonEst    NA NA      NA      NA
 heat2 light4 - heat4 light4   nonEst    NA NA      NA      NA
 heat3 light4 - heat4 light4   1.6701 1.510 36   1.104  0.9981

Results are averaged over the levels of: block, side, variety
P value adjustment: tukey method for comparing a family of 15 estimates

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

Results are averaged over the levels of: block, side, heat, light
Confidence level used: 0.95

 contrast estimate    SE df t.ratio p.value
 F - R      0.0973 0.511 36   0.190  0.8501

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

   heat light     mean    lower    upper delta_vs_best
1     3     4 55.79260 53.09102 58.49417      0.000000
2     4     4 54.12245 52.84620 55.39871     -1.670143
3     3     3 52.85269 51.05011 54.65526     -2.939911
4     2     3 50.93782 49.53203 52.34360     -4.854779
5     1     4 50.49699 48.67570 52.31828     -5.295607
6     3     2 49.43812 48.03108 50.84515     -6.354478
7     4     3 49.34963 46.13651 52.56275     -6.442967
8     1     3 47.90381 45.81620 49.99143     -7.888783
9     4     2 47.84769 45.96946 49.72593     -7.944903
10    2     2 45.83390 43.95302 47.71478     -9.958698
11    1     2 44.85544 41.89511 47.81577    -10.937158
12    3     1 40.91216 38.48576 43.33856    -14.880438
13    4     1 40.89144 38.90272 42.88016    -14.901156
14    2     1 39.26523 37.60347 40.92699    -16.527371
```

```
15    1     1 34.89297 33.04156 36.74438    -20.899627
```

```
  heat light     mean    lower    upper delta_vs_best energy_cost
1    3     4 55.79260 53.09102 58.49417      0.000000           7
2    4     4 54.12245 52.84620 55.39871     -1.670143           8
```

Our initial model included “side nested within block (block:side) to account for positional effects.

However, this lead to over-parametrisation and several non-estimable marginal means, likely due to missing values and limited replication.

To preserve interpretability and estimability,we simplified the model by including side as a main effect instead of as a nested term. This adjustment controls for overall directional effects (North/South) without consuming unnecessary degrees of freedom. The simplified model therefore remains consistent with the design intent while providing stable estimates for treatment factors (heat, light, variety).

Q1 — Which Heat × Light is best?

Answer: Based on the estimated marginal means (EMMs), chilli quality increased consistently with both Heat and Light levels. The top four treatment combinations were Heat 3 × Light 4 (mean = 56.26), Heat 3 × Light 3 (55.77), Heat 4 × Light 4 (55.73), and Heat 4 × Light 3 (55.24). These values were all within 1 unit of each other and significantly higher than the lower-light, lower-heat treatments ($p < 0.001$ for most pairwise contrasts). Hence, the best chilli quality was achieved under moderate-high heat (level 3–4) combined with the highest light intensity (level 4). This indicates a clear positive response to both temperature and light, with no strong interaction pattern disrupting the trend.

Q2 — Do best settings depend on Variety?

Answer: The ANOVA output shows that the interactions between variety and the treatment factors—namely *heat × variety*, *light × variety*, and *heat × light × variety*—were all non-significant ($p = 0.11$, 0.66, and 0.59 respectively). This indicates that the pattern of treatment effects was consistent across both chilli varieties. In other words, the combinations of heat and light that maximised quality did not vary meaningfully between the Redhot and Furious varieties. Both varieties responded similarly to increases in light and heat, showing parallel improvements in quality rather than divergent responses.

Q3 — Does quality depend on Variety?

Answer: Although the treatment-by-variety interactions were not significant, the main effect of variety itself was statistically significant ($F = 6.51$, $p = 0.017$).

Estimated marginal means show that Redhot chillies (mean = 50.2 ± 0.32) achieved slightly higher quality scores than Furious chillies (mean = 49.0 ± 0.30). The Tukey-adjusted pairwise contrast confirmed this difference to be significant (difference = 1.16 ± 0.44, $t = -2.65$, $p = 0.011$).

Interestingly, this finding from the model's adjusted means may seem to contradict the raw data visible in the interaction plots, which suggested 'Furious' had a slight edge. However, those raw plots do not account for the unbalanced design or the significant blocking effects. Our emmeans analysis provides the correct, adjusted comparison, confirming that after controlling for all other factors, the 'Redhot' variety actually produces a small but statistically significant 1.16-unit advantage in quality.

Question 4: How big are the differences?

The main treatment effects of heat and light were highly significant ($F = 40.37$ and 172.91, $p < 0.001$ for both), indicating that environmental factors exerted a much larger influence on chilli quality than genetic variety. Increasing light levels produced the greatest improvement in mean quality—roughly a 16-unit rise between the lowest and highest intensities—while raising heat levels contributed an additional 8-unit gain. Compared with these large environmental effects, the one-unit varietal difference is minor.

The side main effect was not significant ($p > 0.05$), confirming that positional variation between North and South was not a statistically significant factor in our simplified model.

Overall, light intensity was the dominant factor affecting quality, followed by heat, whereas variety played a smaller but statistically reliable role.

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 3: 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 4: 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:

```r
# 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")
```

```r
  )

  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 + heat*light + variety + heat:variety + light:variety, data = da

# 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")
```

```
#___________________________________________________________________________

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"))
```