Problem (circle) : R1 / R2 / A1 / A2 Group number 01 4/12 → 2/6

0.75/1 1. Formatting:

all margins 2.5cm	**informative title**
12 pt size	member names on all pgs
no raw R code or output	all pages numbered
max **7** pages	no blurry plots (**NOT png**)

- too many digits

0.5/1 2. Introduction/Background:

brief statement of scientific question – not correct

all variables defined

3. EDA: not done?

0/2

univariate numerical	bivariate numerical (cor)
univariate graphical	bivariate graphical

0.5/2 4. Model fitting:

define model mathematically – before fitting

state how model fitted (ie, LS)

CLEARLY describe how modelselected – not complete

define all terms – AIC, HSD

0.5/2 5. Model assessment:

CLEARLY state model assumptions: incomplete

1. errors have mean 0
2. errors are homoscedastic (same variance)
3. errors are uncorrelated
4. errors are normally distributed

carry out assessment (graphics):

qq normal plot of residuals, – interpret, don't just conclude

residuals vs. fitted

2.25/8

0.5 / 1

6. Write out final estimated model **mathematically** incomplete + not completely defined (what do the variable terms mean?)

hat on response variable max **2 sig digits** on coefs

0.75 / 1

7. Plots: + number each plot

label size (not too small) informative captions

placement **NOT BLURRY**

0.25 / 1

8. Conclusions

recap analysis state and interpret not correct

main findings

0.25 / 1

9. Overall presentation (clarity of explanations and language, appropriate citations / references):

poor satisfactory good excellent

- project 1 - markdownfile.pdf ?? (p.1)
- incorrect interps (p.1/2/throughout)
- 2-way anova?? (p.3)
- no citations
- what is the purpose of the plot on p.2 ?
- interpret interaction plots, don't just conclude

1.75/4

6.5/12 → 3.25/6

Problem (circle) : R1 / R2 / A1 / A2 Group number 02

1. Formatting:

all margins 2.5cm	**informative title**
12 pt size	member names on all pgs
no raw R code or output	all pages numbered
max **7** pages	no blurry plots (**NOT png**)

1/1

- digits (t/p-values)

2. Introduction/Background:

0.25/1

brief statement of scientific question - not correct

all variables defined Flavor score

3. EDA:

1/2

Also for categorical vars

univariate numerical	bivariate numerical (cor)
univariate graphical	bivariate graphical

4. Model fitting:

1/2

define model mathematically - over-parametrized; i, j, k, l, m, n as written

state how model fitted (ie, LS) not defined

CLEARLY describe how modelselected - which F-tests explain selection procedure not just command

define all terms i/j/.../m, AIC

5. Model assessment:

1.25/2

CLEARLY state model assumptions:

1. errors have mean 0
2. errors are homoscedastic (same variance)
3. errors are uncorrelated - not correct
4. errors are normally distributed

carry out assessment (graphics):

SQUARE qq normal plot of residuals, - interpret plot

residuals vs. fitted - interpret plot

(don't need Shapiro-Wilk, Levene

4.5/8

0.25/1

6. Write out final estimated model **mathematically** where are est. coefs?

hat on response variable max **2 sig digits** on coefs

0.75/1

7. Plots: Fig 1 - density curve not defined

label size (not too small) informative captions

placement **NOT BLURRY**

0.25/1

8. Conclusions

recap analysis state and interpret not correct

main findings

0.75/1

9. Overall presentation (clarity of explanations and language, appropriate citations / references) : tests (?)

poor satisfactory good excellent

- results would be more clear in a table
- interpret interaction plots, don't just conclude
- give p-values (not < 0.01) 0
- anova tables ?

2/4

Problem (circle) : R1 / R2 / A1 / A2 Group number 03

1. Formatting: 1/1

all margins 2.5cm	**informative title**
12 pt size	member names on all pgs
no raw R code or output	all pages numbered
max **7** pages	no blurry plots (**NOT png**)

2. Introduction/Background: 0.5/1 - not only Western meals!!

brief statement of scientific question - imprecise

all variables defined - F (not $\hat{F}$)

3. EDA: 1/2 - write out words

univariate numerical	bivariate numerical (cor)
univariate graphical	bivariate graphical

} incomplete

4. Model fitting: 0.75/2

(*) define model mathematically - not estimated $\hat{F}$, $+\epsilon$ - no residuals as written

state how model fitted (ie, LS)

CLEARLY describe how modelselected - incomplete / somewhat incorrect

define all terms i, j, AIC

- write hyps. mathematically

5. Model assessment: ALL

CLEARLY state model assumptions:

1. errors have mean 0 - resids always have mean 0
2. errors are homoscedastic (same variance)
3. errors are uncorrelated
4. errors are normally distributed

carry out assessment (graphics):

SQUARE qq normal plot of residuals,

residuals vs. fitted - incomplete interpretation

(*) The way you have written the model is incorrect, you need to use indicator variables for factor levels

0.75/1

6. Write out final estimated model **mathematically**

hat on response variable max **2 sig digits** on coefs

The β's are also estimated

0.75/1

7. Plots:

label size (not too small) informative captions

placement **NOT BLURRY**

0.5/1

8. Conclusions

recap analysis (+EDA) state and interpret main findings not completely correct

0.5/1

9. Overall presentation (clarity of explanations and language, appropriate citations / references) : overall logic hard to follow

poor satisfactory good excellent

- refs incomplete
- anova tables?
- why do you use different measures for your different models

5/12 ⇒ 2.5/6

Problem (circle) : R1 / R2 / A1 / A2 Group number 04

1. Formatting:

1/1

all margins 2.5cm	**informative title**
12 pt size	member names on all pgs
no raw R code or output	all pages numbered
max **7** pages	no blurry plots (**NOT png**)

0.25/1 2. Introduction/Background:

- brief statement of scientific question — not clear
- all variables defined — flavor score

1/2 3. EDA:

univariate numerical	bivariate numerical (cor)
univariate graphical	bivariate graphical

↳ Figure 1 perhaps better summarized numerically

4. Model fitting:

0.25/2

- define model mathematically
- state how model fitted (ie, LS)
- CLEARLY describe how modelselected
- define all terms

why multiple 1-way instead of multi-way?

0.5/2 5. Model assessment:

All ↳ interpretation not correct (bottom p.2)

CLEARLY state model assumptions:

1. errors have mean 0
2. errors are homoscedastic (same variance)
3. errors are uncorrelated
4. errors are normally distributed

carry out assessment (graphics):

- qq normal plot of residuals, – SQUARE
- residuals vs. fitted

interpret plots, don't just conclude

3/8

not done?

0/1

6. Write out final estimated model **mathematically**

hat on response variable max **2 sig digits** on coefs

7. Plots:

1/1

label size (not too small) informative captions

placement **NOT BLURRY**

0.25/1

8. Conclusions

recap analysis state and interpret not correct

main findings

0.75/1

9. Overall presentation (clarity of explanations and language, appropriate citations / references) :

poor satisfactory good excellent

- Square QQ plots
- 'normality' (not 'gaussianity')

2/4

0.5/12 → 4.25/6

Problem (circle) : R1 / R2 / A1 / A2 Group number 06

1. Formatting:

1/1

all margins 2.5cm	**informative title**
12 pt size	member names on all pgs
no raw R code or output	all pages numbered
max **7** pages	no blurry plots (**NOT png**)

0.75/1

2. Introduction/Background:

brief statement of scientific question

all variables defined – (flavor score only implicit)

2/2

3. EDA:

univariate numerical	bivariate numerical (cor)
univariate graphical	bivariate graphical

(probably don't need Figure 2)

0.5/2

4. Model fitting:

define model mathematically – why no interactions?

state how model fitted (ie, LS) you fit model (not 'try to')

CLEARLY describe how modelselected

define all terms

errors 'satisfy' (not 'verify') + iid

5. Model assessment: You 'assess', not 'verify', assumptions

1.5/2

CLEARLY state model assumptions:

1. errors have mean 0
2. errors are homoscedastic (same variance)
3. errors are uncorrelated
4. errors are normally distributed

carry out assessment (graphics):

SQUARE – qq normal plot of residuals,

residuals vs. fitted – FULL interpretation

6/8

0.75/1 (define $\hat{y}$)

6. Write out final estimated model **mathematically**

hat on response variable max **2 sig digits** on coefs

0.75/1 → hat\alpha - {1, i}

7. Plots:

label size (not too small) informative captions

placement **NOT BLURRY**

- figure 2 not 'pie charts'

0.25/1 8. Conclusions

recap analysis state and interpret not correct

main findings

0.75/1 9. Overall presentation (clarity of explanations and language, appropriate citations / references) :

poor satisfactory good excellent

→ cite all refs in text (no 'general' refs)

- somewhat superficial

2.5/4

9.25/12 → 4.625/6

Problem (circle) : R1 / R2 / A1 / A2 Group number 08

1. Formatting:

1/1

all margins 2.5cm	**informative title**
12 pt size	member names on all pgs
no raw R code or output	all pages numbered
max **7** pages	no blurry plots (**NOT png**)

1/1

2. Introduction/Background:

brief statement of scientific question

all variables defined

2/2

3. EDA: categorical

univariate numerical bivariate numerical (cor)

univariate graphical bivariate graphical

1.25/2

4. Model fitting:

define model mathematically — over-parametrized as written; ^ is for estimated values; incompletely specified

state how model fitted (ie, LS)

CLEARLY describe how modelselected

define all terms — i, j, k, l, e^r

1.25/2

5. Model assessment: - you assess, not 'verify'

CLEARLY state model assumptions:

1. errors have mean 0
2. errors are homoscedastic (same variance)
3. errors are uncorrelated
4. errors are normally distributed

carry out assessment (graphics):

SQUARE

- qq normal plot of residuals,

residuals vs. fitted — interpret

(don't need Shapiro-Wilk)

6.5/8

- why do you have interactions if they are not significant?

0.5/1

6. Write out final estimated model **mathematically**

hat on response variable max **2 sig digits** on coefs

- no error in est. model

1/1

7. Plots:

label size (not too small) informative captions

placement **NOT BLURRY**

0.5/1

8. Conclusions - use paragraphs be specific

recap analysis state and interpret main findings

0.75/1

9. Overall presentation (clarity of explanations and language, appropriate citations / references) : somewhat incomplete

poor satisfactory good excellent

- don't need Table 1
- Typically you start reducing the model by removing interactions
- integer df

2.75/4

5.75/12 → 2.875/6

Problem (circle) : R1 / R2 / A1 / A2 Group number 10

1. Formatting:

all margins 2.5cm	**informative title**
12 pt size	member names on all pgs
no raw R code or output	all pages numbered
max **7** pages	no blurry plots (**NOT png**)

1/1

2. Introduction/Background:

brief statement of scientific question

all variables defined

1/1

3. EDA: words ??

univariate numerical	bivariate numerical (cor)
univariate graphical	bivariate graphical

1/2 + flavor score

4. Model fitting: no interactions?

define model mathematically - no ↑ with +ε

state how model fitted (ie, LS)

CLEARLY describe how modelselected

define all terms AIC; i

XXXi ??

0.5/2

5. Model assessment:

CLEARLY state model assumptions:

1. errors have mean 0
2. errors are homoscedastic (same variance)
3. errors are uncorrelated
4. errors are normally distributed

0.25/2

carry out assessment (graphics):

SQUARE qq normal plot of residuals,

residuals vs. fitted

interpret these plots

3.75/8

0.75/1

6. Write out final estimated model **mathematically**

hat on response variable max **2 sig digits** on coefs

no error term in estimated model

0.5/1

7. Plots: number all plots

label size (not too small) informative captions

placement **NOT BLURRY**

- what does the plot just above 1.3 mean?

0.25/1

8. Conclusions

recap analysis (incomplete) state and interpret main findings — not correct

0.5/1

9. Overall presentation (clarity of explanations and language, appropriate citations / references) :

poor satisfactory good excellent

- Table 2 p-values un-informative
- just above Table 3: its (not it's)
- 1.9.2 i what is the average flavor score? Without this, cannot tell how good the model fit is
- no references

2/4

Problem (circle) : R1 / R2 / A1 / A2 Group number 11 5.25/12 → 2.625/6

1. Formatting:

1/1

all margins 2.5cm	**informative title**
12 pt size	member names on all pgs
no raw R code or output	all pages numbered
max **7** pages	no blurry plots (**NOT png**)

0.75/1

2. Introduction/Background:

brief statement of scientific question

all variables defined

Flavor Score

0.25/2

3. EDA: - write out words

univariate numerical | bivariate numerical (cor)

univariate graphical | bivariate graphical

0.5/1

4. Model fitting: - over-parametrized as written

define model mathematically – not completely defined (...)

state how model fitted (ie, LS)

CLEARLY describe how modelselected

define all terms i, j, k, l, m

1/2

5. Model assessment: assess, not 'verify'

CLEARLY state model assumptions:

- residuals always sum to 0

1. errors have mean 0
2. errors are homoscedastic (same variance)
3. errors are uncorrelated
4. errors are normally distributed

carry out assessment (graphics):

SQUARE qq normal plot of residuals,

residuals vs. fitted – interpretation not entirely correct

3.5/8

(*) Final model leaves out size, so design/hypotheses not balanced/interpretable

0.25/1

6. Write out final estimated model **mathematically**

hat on response variable max **2 sig digits** on coefs

0.75/1

7. Plots: some plots too large

label size (not too small) informative captions

placement **NOT BLURRY**

not necessarily the entire plot

0.25/1

8. Conclusions

recap analysis state and interpret not correct

main findings

0.5/1

9. Overall presentation (clarity of explanations and language, appropriate citations / references) :

poor satisfactory good excellent

- put refs at end (no footnotes)
- incorrect interpretation of design plot
- use words, your report jumps from bullet point to bullet point

1.75/4

Problem (circle) : R1 / R2 / A1 / A2 Group number 13 9.5/12 → 4.75/6

1. Formatting:

0.75/1

all margins 2.5cm	**informative title**
12 pt size	member names on all pgs
no raw R code or output	all pages numbered
max **7** pages	no blurry plots (**NOT png**)

2. Introduction/Background:

1/1

brief statement of scientific question

all variables defined

2/2

3. EDA:

univariate numerical	bivariate numerical (cor)
univariate graphical	bivariate graphical

1.25/2

4. Model fitting:

over-par. and not linked as written

define model mathematically — 3-way interactions not included

state how model fitted (ie, LS)

CLEARLY describe how modelselected 'incomplete'; stepwise not 'at the same time'

define all terms i, j, k, l, m; AIC

— results for both models

5. Model assessment: — not 'validated'

CLEARLY state model assumptions:

1.25/2

1. errors have mean 0
2. errors are homoscedastic (same variance)
3. errors are uncorrelated — why 'no evidence'
4. errors are normally distributed

carry out assessment (graphics):

SQUARE qq normal plot of residuals,

residuals vs. fitted — incomplete interpretation

(don't need test)

6.25/8

1/1

6. Write out final estimated model **mathematically**

hat on response variable max **2 sig digits** on coefs

1/1

7. Plots:

label size (not too small) informative captions

placement **NOT BLURRY**

0.25/1

8. Conclusions

recap analysis state and interpret main findings – not completely correct

1/1

9. Overall presentation (clarity of explanations and language, appropriate citations / references) :

poor satisfactory good excellent

- square QQ plots

3.25/4

7/12 → 3.5/6

Problem (circle) : R1 / R2 / A1 / A2 Group number 14

0.75/1

1. Formatting:

all margins 2.5cm	**informative title**
12 pt size	member names on all pgs
no raw R code or output	all pages numbered
max **7** pages	no blurry plots (**NOT png**) - (too many digits)

1/1

2. Introduction/Background:

brief statement of scientific question 'are associated with' (not 'influence')

all variables defined

1.25/2

3. EDA:

univariate numerical	bivariate numerical (cor)) incomplete
univariate graphical	bivariate graphical

- Don't need flavor boxplot

1/2

4. Model fitting: over-parametrized as written

define model mathematically

state how model fitted (ie, LS) what does 'Anova framework' mean?

(*) CLEARLY describe how modelselected

define all terms — i, j, k, l, m, ~~AIC~~

- Tests not specified (mathematically)

1/2

5. Model assessment:

CLEARLY state model assumptions:

1. errors have mean 0
2. errors are homoscedastic (same variance)
3. errors are uncorrelated
4. errors are normally distributed

what does 'linearity' mean here? You have specified a linear (in the parameters) model

carry out assessment (graphics):

SQUARE — qq normal plot of residuals, — interpret plot, don't just conclude

residuals vs. fitted — incomplete interpretation

(*) You can specify a 'full' model for stepAIC that is not saturated (look at 'scope')

5/8

0/1 not done?

6. Write out final estimated model **mathematically**

hat on response variable max **2 sig digits** on coefs

1/1

7. Plots:

label size (not too small) informative captions

placement **NOT BLURRY**

0.25/1

8. Conclusions

recap analysis state and interpret not correct

main findings

0.75/1

9. Overall presentation (clarity of explanations and language, appropriate citations / references) : some what incomplete

poor satisfactory good excellent

- you should not have p-values of 0

2/4

7.75/12 → 3.875/6

Problem (circle) : R1 / R2 / A1 / A2 Group number 17

1. Formatting:

1/1

all margins 2.5cm	**informative title**
12 pt size	member names on all pgs
no raw R code or output	all pages numbered
max **7** pages	no blurry plots (**NOT png**)

1/1

2. Introduction/Background:

brief statement of scientific question

all variables defined

1.75/2

3. EDA: 'no sig assoc' = ??

univariate numerical	bivariate numerical (cor)
univariate graphical	bivariate graphical

0.75/2

4. Model fitting: - why do multiple ways? - interpretation + explanation of int. plots

does - parametrized as written

define model mathematically

state how model fitted (ie, LS)

CLEARLY describe how modelselected

define all terms

Tukey HSD (or leave out) - plots not interpreted

1/2

5. Model assessment: model not 'validated', it's assessed

assess, not 'verify'

CLEARLY state model assumptions:

1. errors have mean 0
2. errors are homoscedastic (same variance)
3. errors are uncorrelated
4. errors are normally distributed

carry out assessment (graphics):

SQUARE - qq normal plot of residuals,

residuals vs. fitted

These plots do not 'ensure the validity' of model

explain + interpret plots, don't just conclude

- Don't need Levene's

5.5/8

0.5/1

6. Write out final estimated model **mathematically**

hat on response variable max **2 sig digits** on coefs

1/1

7. Plots:

label size (not too small) (informative captions) QQ normal plot of residuals

placement **NOT BLURRY**

0.25/1

8. Conclusions -not 'confirms'

recap analysis state and interpret not completely correct

main findings

what are the 'valuable insights'?

0.5/1

9. Overall presentation (clarity of explanations and language, appropriate citations / references) :

poor satisfactory good excellent logic somewhat difficult to follow

-cite primary refs

-Conclusions over-stated, also vague

2.25/4

7.25/12 → 3.625/6

Problem (circle) : R1 / R2 / A1 / A2 Group number 18

1. Formatting:

all margins 2.5cm	**informative title**
12 pt size	member names on all pgs
no raw R code or output	all pages numbered
max **7** pages	no blurry plots (**NOT png**)

1/1

0.5/1

2. Introduction/Background:

brief statement of scientific question

all variables defined

1.5/2

3. EDA: summary stats

univariate numerical

bivariate numerical (cor)

univariate graphical

bivariate graphical

Flavor Score Hist

1/2 (*)

4. Model fitting:

unclear; define X's as indicators

define model mathematically (Before fitting + selection which interactions?

state how model fitted (ie, LS)

CLEARLY describe how modelselected stepwise procedure vague

define all terms

('Both directions' is called stepwise)

- size not written in model

0.75/2

5. Model assessment: - assess not 'verify'

CLEARLY state model assumptions:

1. errors have mean 0
2. errors are homoscedastic (same variance) - what are the K pops?
3. errors are uncorrelated
4. errors are normally distributed

carry out assessment (graphics):

qq normal plot of residuals,

residuals vs. fitted

explain plots, then interpret then conclude

(*) no F-value for MSids

4.75/8

0.75/1

6. Write out final estimated model **mathematically**

hat on response variable max **2 sig digits** on coefs

1/1

7. Plots:

label size (not too small) informative captions

placement **NOT BLURRY**

- These results seem to conflict with final model

0.25/1

8. Conclusions

recap analysis state and interpret not correct

main findings

logic hard to follow in parts

0.5/1

9. Overall presentation (clarity of explanations and language, appropriate citations / references):

poor satisfactory good excellent

- refs: how is phylogenetic tree AIC relevant?
- cite primary refs
- your model selection procedure and results are not completely clear

2.5/4

7.25/12 → 3.625/6

Problem (circle) : R1 / R2 / A1 / A2 Group number 19

1/1

1. Formatting:

all margins 2.5cm	**informative title**
12 pt size	member names on all pgs
no raw R code or output	all pages numbered
max **7** pages	no blurry plots (**NOT png**)

0.5/1

2. Introduction/~~Background~~:

brief statement of scientific question

all variables defined

1/2

3. EDA:

univariate numerical	bivariate numerical (cor)
univariate graphical	bivariate graphical

0.75/1

4. Model fitting:

→ 'appropriate' not defined

define model mathematically – incorrectly specified in text; the representation in the appendix is confusing

state how model fitted (ie, LS)

CLEARLY describe how modelselected

define all terms

(*) – final model anova table?

1.5/2

5. Model assessment:

CLEARLY state model assumptions:

1. errors have mean 0
2. errors are homoscedastic (same variance)
3. errors are uncorrelated
4. errors are normally distributed

carry out assessment (graphics):

SQUARE – qq normal plot of residuals,

residuals vs. fitted – interpretation is vague

(*) if you have interactions with size, you should also have size as a main effect, otherwise you no longer have orthogonal effects

4.75/8

0.5/1 (*)

6. Write out final estimated model **mathematically**

hat on response variable max **2 sig digits** on coefs

1/1

7. Plots:

label size (not too small) informative captions

placement **NOT BLURRY**

0.5/1

8. Conclusions

recap analysis state and interpret main findings

0.5/1

9. Overall presentation (clarity of explanations and language, appropriate citations / references) : somewhat incomplete

poor satisfactory good excellent

(*) write the final estimated model as an equation (combine pgs 5+6)

logic somewhat hard to follow

2.5/4

7.5/12 → 3.75/6

Problem (circle) : R1 / R2 / A1 / A2 Group number 20

0.5/1

1. Formatting:

all margins 2.5cm	**informative title**
12 pt size	member names on all pgs
no raw R code or output	all pages numbered
max **7** pages	no blurry plots (**NOT png**)

use more interligne, report is hard to read - too many digits

0.75/1

2. Introduction/Background:

brief statement of scientific question

all variables defined

Flavor score (only implicitly defined)

1.5/2

3. EDA:

univariate numerical	bivariate numerical (cor)
univariate graphical	bivariate graphical

0.75/1 (*)

4. Model fitting:

define model mathematically - why write $\sum_{2}^{3}$??

state how model fitted (ie, LS)

CLEARLY describe how modelselected

define all terms

- what you have is NOT a 'full factorial' model; also Tukey HSD

1.5/2

5. Model assessment: model not 'validated' you assess, not 'verify' you include ints.

CLEARLY state model assumptions:

1. errors have mean 0
2. errors are homoscedastic (same variance)
3. errors are uncorrelated
4. errors are normally distributed

carry out assessment (graphics):

SQUARE - qq normal plot of residuals,

residuals vs. fitted - explain plot (what does 'red line' mean?)

explain clearly

(*) you over-state your findings - analysis does not 'confirm' a causal effect

5/8

(see *) 0.75/1

6. Write out final estimated model **mathematically**

hat on response variable max **2 sig digits** on coefs

1/1

7. Plots:

label size (not too small) informative captions

placement **NOT BLURRY**

0.25/1

8. Conclusions

recap analysis state and interpret main findings not correct

0.5/1

9. Overall presentation (clarity of explanations and language, appropriate citations / references):

poor satisfactory good excellent + Correctness

* re-write Equation (2) to include the numerical values and remove Table 4
- remove excessive space p. 2
- no refs

2.5/4

7/12 → 3.5/6

Problem (circle) : R1 / R2 / A1 / A2 Group number 88

1. Formatting:

1/1

all margins 2.5cm	**informative title**
12 pt size	member names on all pgs
no raw R code or output	all pages numbered
max **7** pages	no blurry plots (**NOT png**)

0.75/1

2. Introduction/Background:

brief statement of scientific question (not quite correct)

all variables defined

1.25/2

3. EDA:

univariate numerical	bivariate numerical (cor) - incomplete
univariate graphical	bivariate graphical

Don't need boxplot

0.75/1

4. Model fitting:

over-parametrized as written

define model mathematically - no $\hat{}$ with $+\epsilon$

state how model fitted (ie, LS)

CLEARLY describe how modelselected

define all terms i, j, k, l, m

- put Table 4 all on 1 page + variable names in Table

5. Model assessment:

1/2

CLEARLY state model assumptions:

1. errors have mean 0
2. errors are homoscedastic (same variance)
3. errors are uncorrelated
4. errors are normally distributed

carry out assessment (graphics):

SQUARE - qq normal plot of residuals,

residuals vs. fitted - explain plot + interpret correctly

4.75/8

0.75/1

y not defined

6. Write out final estimated model **mathematically**

hat on response variable max **2 sig digits** on coefs

0.75/1

7. Plots:

label size (not too small) informative captions

placement **NOT BLURRY**

0.5/1

8. Conclusions

recap analysis (+EDA) state and interpret main findings not correct

0.25/1

9. Overall presentation (clarity of explanations and language, appropriate citations / references) :

logic hard to follow

poor satisfactory good excellent

- Table 6 confusing
- no refs
- why Figure 2 has 2 separate plots?
- Square q-q plots

2.25/4

7.25/12 → 3.625/6

Problem (circle) : R1 / R2 / A1 / A2 Group number 89

0.75/1

1. Formatting:

all margins 2.5cm	**informative title**
12 pt size	member names on all pgs
no raw R code or output	all pages numbered
max **7** pages	no blurry plots (**NOT png**)

too many digits

0.75/1

2. Introduction/Background:

brief statement of scientific question — Don't need (3)

all variables defined — Flavor only implicitly defined

2/2

3. EDA:

univariate numerical	bivariate numerical (cor)
univariate graphical	bivariate graphical

0.75/1 (*) (**)

4. Model fitting: over-parametrized as written

define model mathematically

state how model fitted (ie, LS)

CLEARLY describe how modelselected - Explain any tests mathematically

define all terms i,j,k,l,m, AIC

0.5/2

5. Model assessment: you don't ensure the validity of a model

CLEARLY state model assumptions:

1. errors have mean 0
2. errors are homoscedastic (same variance)
3. errors are uncorrelated
4. errors are normally distributed

carry out assessment (graphics):

SQUARE qq normal plot of residuals, - interpret, don't just conclude

residuals vs. fitted - interpretation not completely correct

explain plots

5/8 (*) (**) How do you deal with the saturated model?
You have 2 size in interaction terms but not as main effect
- this means the effects are not orthogonal !!

0/1 not done?

6. Write out final estimated model **mathematically**

(**hat** on response variable max **2 sig digits** on coefs

7. Plots:

1/1

label size (not too small) informative captions

placement **NOT BLURRY**

0.5/1

8. Conclusions

recap analysis state and interpret not correct

+EDA main findings

0.75/1

9. Overall presentation (logic unclear in parts

(clarity of explanations and language, appropriate citations / references) :

poor satisfactory good excellent

- square QQ plots
- cite primary refs (not [5])
- anova tables would be much clearer if you use words rather than symbols

2.25/4

5.75/12 → 2.875/6

Problem (circle) : R1 / R2 / A1 / A2 Group number 30

1. Formatting: 0.5/1

all margins 2.5cm

informative title

12 pt size

member names on all pgs

no raw R code or output — Top p.7

all pages numbered

max **7** pages

no blurry plots (**NOT png**)

2. Introduction/Background: 0.75/1

brief statement of scientific question (not entirely correct)

all variables defined

3. EDA: 1.25/2 write out words. Figure 3 unclear because only 2 levels of G + H (??)

univariate numerical

bivariate numerical (cor)

univariate graphical

bivariate graphical

Don't need figure 2

4. Model fitting: 0.25/2 Don't need F formula. randomization description not entirely correct

define model mathematically — specific to THIS problem (not eq. 1)

state how model fitted (ie, LS)

CLEARLY describe how modelselected (first you say Eq 4 is 'final' model, the you put int; size SS not = 0)

define all terms

5. Model assessment: 0.5/2 assess not 'verify'

CLEARLY state ALL model assumptions:

1. errors have mean 0
2. errors are homoscedastic (same variance) — Here, the 'groups' are the cross classified obs, NOT separately 'for each var.'
3. errors are uncorrelated
4. errors are normally distributed

carry out assessment (graphics):

SQUARE qq normal plot of residuals,

residuals vs. fitted

+ interpret plots

Figure 3: Grow 1 2 3 4; Hold 1 2 3 4 — 2 2

not G 1 2 / 3 4, H 1 2 / 3 4

3.25/8

which is Final model?

1/1

6. Write out final estimated model **mathematically**

hat on response variable max **2 sig digits** on coefs

0.25/1

7. Plots:

label size (not too small) informative captions

placement **NOT BLURRY**

0.25/1

8. Conclusions

recap analysis state and interpret main findings

not completely correct

+ completeness/correctness

0.5/1

9. Overall presentation (clarity of explanations and language, appropriate citations / references): somewhat insufficient

poor satisfactory good excellent

- Table 3 differs slightly from Eq. 4
- Don't need both Tables 3+4, just put equation with final selected model

2.5/4