6.25/12 → 3.125/6

Problem (circle) : R1 / R2 / A1 / A2 Group number 44

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**)

0.75/1 - too many digits

2. Introduction/Background:

brief statement of scientific question — not entirely correct

all variables defined

1.25/2 3. EDA:

univariate numerical — bivariate numerical (cor)

univariate graphical — bivariate graphical — all pairs

- Don't need boxplots

1.25/2 4. Model fitting: Don't need single predictor models

define model mathematically — 'seems to be' ??

state how model fitted (ie, LS)

CLEARLY describe how modelselected — - not R^2, always goes up with more vars

define all terms

'statistic' ? (Table 1)

- Table not 'tab'

0.5/1 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):

qq normal plot of residuals,

residuals vs. fitted — 'good plot'? - explain plots, then interpret, then conclude

(not Fitted vs Resids

⇒ you plot Y vs X

4.5/8

0.5/1

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

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

7. Plots:

+ number all figs

0.5/1

label size (not too small) informative captions

placement **NOT BLURRY**

Bivariate plot shapes bizarre

0.25/1

8. Conclusions

recap analysis state and interpret not completely correct

main findings

0.5/1

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

logic sometimes hard to follow

poor satisfactory good excellent

- Appendix table 1 incomplete, also p=0?

1.75/2

7/12 → 3.5/6

Problem (circle) : R1 / R2 / A1 / A2 Group number 45

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

0.75/1

2. Introduction/Background:

brief statement of scientific question – make more clear

all variables defined ('response' variable (not 'target'))

1/2

3. EDA: incomplete

univariate numerical — bivariate numerical (cor) ?

univariate graphical — bivariate graphical

Don't need box/dot plots — all pairs

1.25/2

4. Model fitting: not 'verify'

$2^{+\epsilon}$ – define model mathematically Don't need individual predictors

state how model fitted (ie, LS) least squares

CLEARLY describe how modelselected

define all terms

where is R^2 / R^2_{adj}

5. Model assessment: assess not 'verify'

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):

qq normal plot of residuals,

residuals vs. fitted

explain plots, then interpret then conclude

⇒ how do these plots aid in assessing assumptions?

5/8

0.75/1

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

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

0.5/1

7. Plots:

label size (not too small) informative captions (labels)

placement **NOT BLURRY**

Figure 6 too big, and why do you have it?

0.25/1

8. Conclusions

recap analysis state and interpret main findings

0.5/1

9. Overall presentation (clarity of explanations and language, appropriate citations / references) : complete spell check + grammar check

poor satisfactory good excellent

- what is the 'obtained R2'
- Don't need to give R commands

2/4

7.75/12 → 3.875/6

Problem (circle) : R1 / R2 / A1 / A2 Group number 61

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**)

model descriptions

2. Introduction/Background:

1/1

brief statement of scientific question

all variables defined

1.25/2

3. EDA:

univariate numerical	bivariate numerical (cor)
univariate graphical	bivariate graphical

combine Figs 2+3

histograms not boxplots

square

0.75/2

4. Model fitting:

define model mathematically

state how model fitted (ie, LS)

CLEARLY describe how modelselected

define all terms AIC

– why summarize so many model results

5. Model assessment:

assess, not 'validate'

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 – incorrect plot) explain then interpret then conclude

4.75/8

1/1

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

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

4/1

7. Plots:

label size (not too small) informative captions

placement **NOT BLURRY**

0.5/1

8. Conclusions

recap analysis state and interpret not correct

main findings – correlation does not need the model.

0.5/1

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

spell check 'appears' not sig ? Either it is or not

poor satisfactory good excellent

-no refs

3/4

8/12 → 14/16

Problem (circle) : R1 / R2 / A1 / A2 Group number 62

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**)

- too many digits

0.75/1

2. Introduction/Background:

brief statement of scientific question - 'estimate' not 'understand'

all variables defined (crime only implicit)

1.5/2

3. EDA:

univariate numerical	bivariate numerical (cor)
univariate graphical	bivariate graphical all pairs

Don't need boxplots

1.25/2

4. Model fitting:

define model ~~mathematically~~ Before estimated

state how model fitted (ie, LS)

CLEARLY describe how modelselected

define all terms

- put results in a table, not inline

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
4. errors are normally distributed

carry out assessment (graphics):

qq normal plot of residuals,

residuals vs. fitted

explain plots, then interpret then conclude

Incomplete

5.5/8

1/1

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**

0.25/1

8. Conclusions

recap analysis state and interpret

main findings

- model not necessary to explain 1+2

0.5/1

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

poor satisfactory good excellent

- use sentences, not bullet points

2.5/4

7.75/12 → 3.875/6

Problem (circle) : R1 / R2 / A1 / A2 Group number 63

1. Formatting:

0.5/1

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

model specifications — Too many digits

0.75/1 2. Introduction/Background:

- brief statement of scientific question — not entirely correct
- all variables defined — variables, not 'factors'

1.75/1 3. EDA:

univariate numerical	bivariate numerical (cor)
univariate graphical	bivariate graphical

could combine To make more compact

Don't need Table 2

1.25/2 4. Model fitting: LS not a 'model' — incorrect interpretation of coefs

- define model mathematically
- state how model fitted (ie, LS)
- CLEARLY describe how modelselected
- define all terms — AIC 'much improvement'?

1.25/2 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 (not just funnel)

explain plots then interpret then conclude

5.5/8

which model? Don't need more than 1

0.75/1 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**

you could include in text if you remove some extraneous stuff

0.25/1 8. Conclusions

recap analysis state and interpret main findings not clear

several vague statement

0.5/1 9. Overall presentation (clarity of explanations and language, appropriate citations / references):

you could make your logic clearer

poor satisfactory good excellent

2.25/4

Problem (circle) : R1 / R2 / A1 / A2 Group number 64 6.75/12 → 3.375/6

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**)

0.5/1

2. Introduction/Background:

brief statement of scientific question imprecise

all variables defined - not clear, Define in text before Table

1.25/2 3. EDA: ↳ you don't need to re-iterate numerical results in text use more paragraphing

univariate numerical	bivariate numerical (cor)
univariate graphical	bivariate graphical

histograms square QQ plots - Figure 3 redundant ↳ square

1/2 4. Model fitting: 'seems that'? High stat sig of coefs (not vars)

define model mathematically no ^, +ε

state how model fitted (ie, LS)

CLEARLY describe how modelselected - incorrect coef interp (worship)

define all terms

0.75/2 5. Model assessment: - why Figure 5? - you have a number of vague statements assessed not 'verified'

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): (clearly) Explain plots then interpret then Conclude

SQUARE qq normal plot of residuals,

residuals vs. fitted

4.25/8

1/1

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**

0.25/1

8. Conclusions

recap analysis state and interpret

main findings - not correct you should indicate what you find from the model, not correlations

0.5/1

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

poor satisfactory good excellent

logic not completely clear

2.5/4

8.5/12 → 4.25/6

Problem (circle) : R1 / R2 / A1 / A2 Group number 65

1. Formatting:

0.75/1

all margins 2.5cm — **informative title**

12 pt size Throughout — 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

2. Introduction/Background:

1/1

brief statement of scientific question

all variables defined

3. EDA:

1.25/2

univariate numerical — bivariate numerical (cor)

univariate graphical — bivariate graphical — square

- you don't need plots to know direction of correlation, just look at the number

4. Model fitting: (not 'interact')

1.5/2

define model mathematically

state how model fitted (ie, LS) squares

CLEARLY describe how modelselected

define all terms AIC

p<0.006

- Don't need results from all models

5. Model assessment:

1.75/2

CLEARLY state model assumptions:

1. errors have mean 0 - resids always sum to 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,

residuals vs. fitted

6.25/8

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:

label size (not too small) informative captions

placement **NOT BLURRY**

0.5/2

8. Conclusions not 'statistically proven'

recap analysis state and interpret main findings

0.75/1

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

poor satisfactory good excellent

2.25/4

6.5/12 → 3.25/6

Problem (circle) : R1 / R2 / A1 / A2 Group number 66

1. Formatting:

0.75/1

all margins 2.5cm — **informative title**

12 pt size + more interline — 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

brief statement of scientific question - not clear

all variables defined

3. EDA:

1/2

univariate numerical (all summary stats) — bivariate numerical (cor)

univariate graphical — bivariate graphical (all pairs)

Don't need boxplots

4. Model fitting:

1/2

define model mathematically (variable, not 'factor')

state how model fitted (ie, LS)

CLEARLY describe how modelselected

define all terms (variable importance)

*

Figure 5: you can get the same info from the correlations

5. Model assessment:

0.75/2

First Explain plots, then interpret then conclude

CLEARLY state model assumptions:

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

(approx in parentheses — for valid F-tests you need normality)

carry out assessment (graphics):

SQUARE - qq normal plot of residuals,

residuals vs. fitted — incomplete interpretation

pred | ← square plot

obs

4/8

* what does 'model is stat sig' mean?

1/1

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**

0.25/1

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

- There is a lot of extraneous material

2.5/4

9.5/12 → 4.75/6

Problem (circle) : R1 / R2 / A1 / A2 Group number 67

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**)

too many digits

1/1

2. Introduction/Background:

brief statement of scientific question

all variables defined

1.25/2

3. EDA:

all summary stats

univariate numerical	bivariate numerical (cor)
univariate graphical	bivariate graphical all pairs

1.5/2

4. Model fitting: p-value 0?

define model mathematically

state how model fitted (ie, LS)

CLEARLY describe how modelselected not just proposed

define all terms

- interpret results
- Don't need 'Sig of Coefs' section

5. Model assessment:

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, - explain plots before interpreting and concluding

residuals vs. fitted - incomplete interpretation

6/8

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**

0.75/1

8. Conclusions

recap analysis +EDA state and interpret

main findings

0.75/1

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

poor satisfactory good excellent

3.5/4

7/12 → 3.5/6

Problem (circle) : R1 / R2 / A1 / A2 Group number 68

1. Formatting:

0.75/1

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

0.75/1

2. Introduction/Background:

- brief statement of scientific question — not entirely correct
- all variables defined

0.75/2

3. EDA:

- univariate numerical
- bivariate numerical (cor)
- univariate graphical — histograms
- bivariate graphical — Don't need plot to determine correlation

1.25/2

4. Model fitting: Don't need all univariate results

- define model mathematically — no ^ with tc
- state how model fitted (ie, LS) — BEFORE fitting
- CLEARLY describe how modelselected
- define all terms

1.25/2

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

} First explain plots, then interpret then conclude

4.75/8

0.5/1

6. Write out final estimated model **mathematically** — not ∈

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

0.75/1

7. Plots:

label size (not too small) — informative captions — plots don't 'analyze'

placement — **NOT BLURRY**

0.25/1

8. Conclusions

recap analysis — state and interpret main findings — not completely correct

0.75/1

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

poor — satisfactory — good — excellent

- How did you select k in k-means? Explain the algorithm or delete this part

2.25/4

6.75/12 → 3.325/6

Problem (circle) : R1 / R2 / A1 / A2 Group number 70

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 - Do you study sociological reasons? - What does 'linked to' mean

all variables defined → limit scope to the specific statistical question(s) you study

1.25/2

3. EDA:

univariate numerical	bivariate numerical (cor)
univariate graphical	bivariate graphical

histograms not boxplots

combine figures 2+3

combine figure 4 with figure 3 (square)

not 'influence'

4. Model fitting:

0.75/2

define model mathematically no ^ with +e

state how model fitted (ie, LS)

CLEARLY describe how modelselected

define all terms (autocorrelation) AIC, BIC

0.5/2

5. Model assessment:

CLEARLY state model assumptions:

not residuals

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 then interpret plots, then conclude

- SW / DW tests not described
- GLS model not defined

4.25/8

Estimated or fitted model (not 'adjustement')

0.75/1

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

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

including autocor est?

0.75/1

7. Plots:

label size (not too small) informative captions

placement **NOT BLURRY**

0.5/1

8. Conclusions many vague statements

recap analysis not 'confirms relationships'

state and interpret (Expected crime rate) not correct

main findings

0.5/1

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

poor satisfactory good excellent

- write out Figure (not Fig.) in text
- Regional: not 'adjustment'; 'influencing'?
- Northern + York are geographically close as well
- somewhat superficial and incomplete

2.5/4

8/12 → 4/6

Problem (circle) : R1 / R2 / A1 / A2 Group number 71

1. Formatting:

0.75/1

all margins 2.5cm

12 pt size

no raw R code or output

max **7** pages

informative title (Don't need R2)

member names on all pgs

all pages numbered

no blurry plots (**NOT png**)

- too many digits

2. Introduction/Background:

1/1

brief statement of scientific question

all variables defined (not 'factors')

3. EDA: put all summary stats in a Table, unorthodox representation

1.25/2

univariate numerical

univariate graphical

bivariate numerical (cor)

bivariate graphical all pairs

- Don't need Figure 3

1.5/2

4. Model fitting:

define model mathematically - no $\hat{}$, +6

state how model fitted (ie, LS)

CLEARLY describe how modelselected

define all terms

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

not approx: this is needed for F-test validity

carry out assessment (graphics):

SQUARE - qq normal plot of residuals,

residuals vs. fitted

Explain then interpret then conclude

5.5/2

1/1

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**

Figure 27

0.25/1

8. Conclusions

recap analysis state and interpret

what does 'most consistent' mean

not correct

main findings

- not why only corr. and not causal - the data are observational

0.5/1

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

poor satisfactory good excellent

- incorrect interpretation of R^2_{adj}

2.5/4

6.5/12 → 3.25/6

Problem (circle) : R1 / R2 / A1 / A2 Group number 72

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**)

- too many digits

0.75/1 2. Introduction/Background:

brief statement of scientific question - not 'correlate with' — you are building a model for prediction (or explanation)

all variables defined

1/2 3. EDA: all summary stats

univariate numerical	bivariate numerical (cor) all
univariate graphical	bivariate graphical all pairs

- Do not summarize underneath plots — hard to read

1.25/2 4. Model fitting:

define model mathematically no ^, +ε

state how model fitted (ie, LS)

CLEARLY describe how modelselected

define all terms

5. Model assessment:

1/2

CLEARLY state model assumptions:

not resids

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

not resids - used for assessment. resids always sum to 0

carry out assessment (graphics):

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

explain plots then interpret then conclude

assess these assumptions

4.25/8

1/1

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

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

0.5/1

7. Plots:

+ number all figures

label size (not too small) informative captions

placement **NOT BLURRY**

0.25/1

8. Conclusions

recap analysis state and interpret not correct

main findings

0.5/1

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

poor satisfactory good excellent

2.25/4

8.75/12 → 4.375/6

Problem (circle) : R1 / R2 / A1 / A2 Group number 73

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 entirely correct

all variables defined

2/2

3. EDA:

univariate numerical bivariate numerical (cor)

univariate graphical bivariate graphical

Don't need Figure 3

1.5/2

4. Model fitting: Don't use the term 'significantly' in the colloquial sense, only in the context of a statistical test

define model mathematically

state how model fitted (ie, LS)

CLEARLY describe how modelselected

define all terms partial correlation, AIC

incorrect interpretation of R^2_{adj}

0.75/2

5. Model assessment: use more paragraphing

CLEARLY state model assumptions: - on errors, assessed by examining resids

not resids

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

carry out assessment (graphics): - Don't interpret plots in caption

qq normal plot of residuals, } - First explain, then interpret then conclude

residuals vs. fitted

- SW / DW not defined
- These tests don't 'confirm'

6/8

0.5/1 (where is this - seems to just be implicit)

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 (ok) state and interpret not correct

main findings

1/1

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

poor satisfactory good excellent

2.75/4

6.75/12 → 3.375/6

Problem (circle) : R1 / R2 / A1 / A2 Group number 74

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**)

- too many digits (p-values p. 4)

1/1 2. Introduction/Background:

brief statement of scientific question

all variables defined (not 'factors')

0.75/2 3. EDA:

univariate numerical — bivariate numerical (cor)

univariate graphical — bivariate graphical all pairs

Don't need to xplot

1/2 4. Model fitting:

define model mathematically

state how model fitted (ie, LS)

CLEARLY describe how modelselected

define all terms (β's, AIC)

what does 'robust' mean here?

5. Model assessment:

0.5/2 not residuals

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,

residuals vs. fitted

First explain, then interpret then conclude

1 incorrect interpretation

4/8

↑ text (= not ≈)

1/1

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 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

- no refs

2.75/4

7.75/12 → 3.875/6

Problem (circle) : R1 / R2 / A1 / A2 Group number 75

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 entirely precise

all variables defined

1.75/2 3. EDA: (cont'd) add loess curves to pairs plots and remove loess graphs

univariate numerical — bivariate numerical (cor)

univariate graphical — bivariate graphical

Don't need boxplots — square

1.25/2 * 4. Model fitting: why single variable results? (Before results)

define model mathematically

state how model fitted (ie, LS) — Before fitting

~~CLEARLY~~ describe how model selected

define all terms AIC, BIC

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

not resids

carry out assessment (graphics):

SQUARE — qq normal plot of residuals,

residuals vs. fitted

First explain plots, then interpret

* A hyp test DOES NOT DETERMINE which hyp is true - you only make a decision (reject or not)

5.25

1/1

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

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

0.75/1

7. Plots:

+ number all figures

label size (not too small) informative captions

placement **NOT BLURRY**

0.25/1

8. Conclusions

vague; also not all linked to model

recap analysis state and interpret

main findings

0.5/1

logic somewhat hard to follow

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

poor satisfactory good excellent

- no refs

2.5/4

7.5/12 → 3.75/6

Problem (circle) : R1 / R2 / A1 / A2 Group number 76

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**)

↳ re-arrange and remove extraneous material

0.75/1 2. Introduction/Background: Don't state results in intro

brief statement of scientific question – make concise, it's hard to follow

all variables defined

1.5/2 3. EDA: – out of order

univariate numerical	bivariate numerical (cor)
univariate graphical	bivariate graphical all pairs

1.5/2 4. Model fitting:

define model mathematically

state how model fitted (ie, LS)

CLEARLY describe how modelselected

define all terms

- incorrect interpretation of R^2_{adj}

0.75/1 5. Model assessment: don't 'ensure the validity' assess, don't validate

CLEARLY state model assumptions: – assumptions on errors, assessment using resids

not resids –

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

is ϵ = mean resids? This's always = 0

carry out assessment (graphics):

qq normal plot of residuals, I don't see it

residuals vs. fitted – incomplete interpretation

- RMSE, not MAE; and need to compare it to Y

5.25/6 – DW not described + no p-values

0.75/1

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

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

0.5/1

7. Plots:

label size (not too small) informative captions

placement **NOT BLURRY**

Don't need excessive reasoning, just description

(somewhat confusing 'may be'?

0.5/1

8. Conclusions – vague

recap analysis (only implicit) state and interpret – not correct

main findings – 'suggests'?

logic is hard to follow

0.5/1

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

poor satisfactory good excellent

- use 1-column format
- no refs
- many extraneous/run-on sentences

2.25/4

6.25/12 → 3.125/6

Problem (circle) : R1 / R2 / A1 / A2 Group number 77

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	all pages numbered
max **7** pages	no blurry plots (**NOT png**)

model specifications — too many digits

1/1

2. Introduction/Background:

brief statement of scientific question

all variables defined

1.75/2

3. EDA:

univariate numerical	bivariate numerical (cor)
univariate graphical	bivariate graphical

Don't need boxplots

0.25/2

4. Model fitting:

define model mathematically — not R formulas

state how model fitted (ie, LS)

CLEARLY describe how modelselected

define all terms AIC (you minimize not max)

0.5/2

5. Model assessment: 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

Explain plots, then interpret, then conclude

4/8

0.75/1

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**

0.25/1

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

- no refs
- very superficial - you need to explain why you are doing what you are doing

2.25/4

7.5/12 → 3.75/6

Problem (circle) : R1 / R2 / A1 / A2 Group number 78

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: Table with all summary stats Combine 4a + 4b

1.25/2

univariate numerical — bivariate numerical (cor) — square

univariate graphical — bivariate graphical ← Figure 3 redundant

histograms not boxplots

4. Model fitting:

(*)

define model mathematically

state how model fitted (ie, LS)

CLEARLY describe how modelselected

define all terms : X also includes intercept, not just predictors

1.25/2

- Compare RMSE to $\bar{Y}$

AIC / BIC

5. Model assessment:

Put text before plots

0.75/2

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 plots, then interpret then conclude

(*) hypothesis test for F-stat? Also, since different df you can only compare F-stats by p-value

5.25/8

0.5/1 (only implied)

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

grammar check

0.5/1

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

poor satisfactory good excellent

- no refs

2.25/4

7.25/12 → 3.625/6

Problem (circle) : R1 / R2 / A1 / A2 Group number 79

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**)

1/1

2. Introduction/Background:

brief statement of scientific question

all variables defined

1.5/2

3. EDA:

univariate numerical	bivariate numerical (cor)
univariate graphical	bivariate graphical

combine figures 2a/b/c

- histograms not boxplots

1/2

4. Model fitting: (pretty much always keep intercept)

define model mathematically

state how model fitted (ie, LS)

CLEARLY describe how modelselected R^2 or R^2_{adj}

define all terms

- put results in a table

use paragraphs

5. Model assessment:

0.75/2

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): – Explain plots then interpret then conclude

SQUARE- qq normal plot of residuals,

residuals vs. fitted interpretation not completely correct

5/8

1/1

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

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

0.5/1

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

0.5/1

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

poor satisfactory good excellent

- use more paragraphing
- somewhat superficial + incomplete

2.25/4

9/12 → 4.5/6

Problem (circle) : R1 / R2 / A1 / A2 Group number 80

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

1.5/2

3. EDA:

univariate numerical — bivariate numerical (cor)

univariate graphical — bivariate graphical

1.25/2

4. Model fitting:

define model mathematically

state how model fitted (ie, LS)

CLEARLY describe how modelselected

define all terms

- incorrect coef interps - interpretation of R^2_{adj}
- R^2 = 0.39 or 0.319?

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

carry out assessment (graphics):

SQUARE - qq normal plot of residuals,

residuals vs. fitted

First explain plots, then interpret then conclude

- Don't use the term 'significant' outside context of a statistical test

5.75/8

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**

0.5/1

8. Conclusions

vague

recap analysis (+EDA) state and interpret main findings

0.75/1

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

poor satisfactory good excellent

3.25/4

Problem (circle) : R1 / R2 / A1 / A2 Group number 81 8/12 → 4/6

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**)

- too many digits

0.75/1 2. Introduction/Background:

brief statement of scientific question - not entirely precise

all variables defined

1/2 3. EDA: Table with all summary stats — unorthodox representation

univariate numerical — bivariate numerical (cor) - not informative & include both; worship/grim not exactly

univariate graphical — bivariate graphical all pairs

- histograms not boxplots

1.25 4. Model fitting: RMSE (and compare to $\bar{Y}$)

define model mathematically

R^2 or Radj state how model fitted (ie, LS)

CLEARLY describe how modelselected - you have a vague statement

define all terms AIC

- penalized regression methods not typical in very low dimensional settings (like here)

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):

qq normal plot of residuals, — Explain plots before interpreting

residuals vs. fitted

5.25/8

1/1

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 use paragraphing

(+EDA) R^2_{adj} misinterpreted not correct

recap analysis state and interpret

(variables not 'factors') main findings

0.5/1

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

poor satisfactory good excellent

- many vague statements + sweeping generalities

2.75/4

Problem (circle) : R1 / R2 / A1 / A2 Group number 82 8.25/12 → 4.125/6

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

0.75/1 2. Introduction/Background:

brief statement of scientific question not entirely correct

all variables defined

1.5/2 3. EDA:

why is figure 2 so big? unorthodox representation; cor crime/church=0?

univariate numerical	bivariate numerical (cor)
univariate graphical	bivariate graphical

square combine these

1.25/2 4. Model fitting:

define model mathematically no ^, +ε

state how model fitted (ie, LS)

CLEARLY describe how modelselected somewhat imprecise description

define all terms

- summarize info from paragraph bottom p.5 in a Table

5. Model assessment:

CLEARLY state model assumptions:

1/2

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

carry out assessment (graphics): diagnostic plots

SQUARE—qq normal plot of residuals,

residuals vs. fitted

explain plots then interpret then conclude

5.5/8

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

1/1

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 - not logistic

recap analysis state and interpret not correct

main findings

0.5/1

9. Overall presentation (clarity of explanations and language, appropriate citations / references) : logic sometimes difficult to follow grammar check

poor satisfactory good excellent

2.75/4

7.75/12 → 3.875/6

Problem (circle) : R1 / R2 / A1 / A2 Group number 83

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**)

- too many digits

2. Introduction/Background: 0.75/1

brief statement of scientific question — not entirely precise

all variables defined

3. EDA: 1.25/2

univariate numerical — bivariate numerical (cor) — hard to read

univariate graphical — bivariate graphical — SQUARE

Don't need figure 2

- SQUARE QQ

(interpretation all vars. + not 'matrice de dispersion')

4. Model fitting: 1/2

define model mathematically

state how model fitted (ie, LS)

CLEARLY describe how modelselected

define all terms — AIC — put results in a table not inline

$Y = \beta_0 + \dots + \epsilon$ [not $Y = 172.894\dots$]

5. Model assessment: 1.5/2 — évaluer pas vérifier

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,

residuals vs. fitted — interpretation not entirely correct

5.25/8

0.75/1 Hard to read this equation

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

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

no ϵ

0.5/1

7. Plots:

label size (not too small) informative captions

placement **NOT BLURRY**

0.5/1

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

2.5/4

Problem (circle) : R1 / R2 / A1 / A2 Group number 84 8/12 → 4/6

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 — all pages numbered

max **7** pages — no blurry plots (**NOT png**)

middle p.5 — Tables+Figures

1/1 2. Introduction/Background:

brief statement of scientific question

all variables defined — not factors

1.5 3. EDA:

univariate numerical — bivariate numerical (cor)

univariate graphical — bivariate graphical – square

combine these into 1 graph

crime rate vs. (other vars)

- Boxplots depict not 'examine'

1.25/2 4. Model fitting:

define model mathematically

state how model fitted (ie, LS)

CLEARLY describe how modelselected

define all terms

- write out hyp tests mathematically

1.25/2 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, – plot doesn't check 'suggesting'?

residuals vs. fitted — incomplete; also, the plot does not evaluate assumptions, you use the plot to evaluate

5.5/8

1/1

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

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

0.5/1

7. Plots:

label size (not too small) informative captions

placement **NOT BLURRY**

0.5/1

8. Conclusions

recap analysis state and interpret main findings not completely correct

0.5/1

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

poor satisfactory good excellent

- layout sometimes looks off
- cite primary sources

2.5/4

9.25/12 → 4.625/6

Problem (circle) : R1 / R2 / A1 / A2 Group number 85

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

(not 'factors')

3. EDA:

univariate numerical bivariate numerical (cor)

univariate graphical bivariate graphical

1.5/2

histograms

4. Model fitting:

define model mathematically no $\hat{}$, $+\epsilon$

state how model fitted (ie, LS)

CLEARLY describe how modelselected

define all terms

1.5/2

R^2_{adj} drops from 7% from model 2 to 3 = ??

5. Model assessment: not 'verify'

0.75/2

CLEARLY state model assumptions:

not resids

1. errors have mean 0 sum of resids always = 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

SW not defined

First explain the plots then interpret then conclude

5.75/8

1/1

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

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

(only use scientific notation if really necessary)

1/1

7. Plots:

label size (not too small) informative captions

placement **NOT BLURRY**

0.75/1

8. Conclusions

recap analysis (+EDA) state and interpret main findings

0.75/1

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

poor satisfactory good excellent

- Section (#) (not section ##)
- square QQ plots

3.5/4

Problem (circle) : R1 / R2 / A1 / A2 Group number 97 7.5/12 → 3.75/6

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**)

0.75/1 2. Introduction/Background:

brief statement of scientific question imprecise

all variables defined not 'factors'

1/2 3. EDA: all summary stats unorthodox representation

univariate numerical, bivariate numerical (cor)

univariate graphical bivariate graphical – all pairs) combine

histograms for all variables Figure 3 displays (not 'explores')

1.5/2 4. Model fitting:

define model mathematically

state how model fitted (ie, LS)

CLEARLY describe how modelselected

define all terms Mallows, BIC

R^2_{adj} interpretation incorrect

0.75/2 5. Model assessment: assess not 'validate', not 'ensure';

CLEARLY state model assumptions: also it's assumptions on the errors (not resids)

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

carry out assessment (graphics): The plots don't evaluate, you use them to evaluate

SQUARE – qq normal plot of residuals,

residuals vs. fitted incomplete interpretation

Explain plots then interpret then conclude

4.75/8

1/1

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

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

0.5/1

7. Plots:

label size (not too small) informative captions

placement **NOT BLURRY**

0.5/1

8. Conclusions

recap analysis state and interpret main findings

(linear assoc)

'suggests'?

not completely correct

(+EDA)

0.75/1

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

poor satisfactory good excellent

2.75/4