GLM: Name NB 9.75/12 → 4.875/6

1. Formatting:

0.75/0.75

- all margins 2.5cm
- 12 pt size
- **no raw R code or output**
- max **10** pages
- informative title
- name on all pages
- all pages numbered
- no blurry plots (**NOT png**)

(too many digits)

1/1

2. Introduction/Background:

- brief statement of scientific question
- all variables defined

3. EDA:

1.5/2

- univariate numerical → (5-number summary)
- univariate graphical
- bivariate numerical (cor)
- bivariate graphical → all pairs

Fig 4 unnecessary + explanation

→ Don't need boxplots (Fig 2)

4. Model fitting:

1.25/2

- give mathematical definition of model
- state how model fitted (ie, maximum likelihood) – Don't need numerical details
- CLEARLY describe how model selected
- define all terms

→ what vars are you using?

5. Model assessment:

- CLEARLY state model assumptions:

1.5/2

1. count outcome Poisson
2. independent obs
3. linear relation between log count and linear predictor
4. conditional mean = conditional variance – ?? 2 numbers are always proportional

- carry out assessment (numerical / graphics):
  scatterplots (linearity assumption)

log-odds?

* Sec. 2.1 goes in Model fitting

6/7.75

1/1 (write coefs inline)

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

hat on response variable (ok if coefs in table)

max **2 sig digits** on coefs

1/1.25 7. Plots: (shapes) + incomplete

label size (not too small)

captions

placement

NOT BLURRY

0.5/1 8. Conclusions

(+EDA

recap analysis

state main findings

somewhat vague at end 'seems to be'?

1.25/1.25 9. Overall presentation (clarity of explanations, appropriate citations / references) :

poor satisfactory good excellent

10. Other comments:

3.75/4.25

GLM: Name LB 5.5/12 → 2.75/6

1. Formatting:

0.5/0.75

all margins 2.5cm

12 pt size

no raw R code or output

max **10** pages

informative title

name on all pages

all pages numbered

no blurry plots (**NOT png**)

0.5/1

2. Introduction/Background: use your own words

brief statement of scientific question

all variables defined

0.25/2

3. EDA:

univariate numerical

univariate graphical

bivariate numerical (cor)

bivariate graphical

all pairs

1/2

4. Model fitting:

give mathematical definition of model

state how model fitted (ie, maximum likelihood)

CLEARLY describe how model selected

define all terms

- if you have interaction you should normally keep lower order terms

0.75/2

5. Model assessment: you assess not verify

CLEARLY state model assumptions: incomplete

1. count outcome Poisson
2. independent obs
3. linear relation between log count and linear predictor
4. conditional mean = conditional variance

carry out assessment (numerical / graphics):
scatterplots (linearity assumption)

SQUARE QQ

p.4 (top) very vague

3/7.75

0.25/1 write numerical equation

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

hat on response variable (ok if coefs in table) max **2 sig digits** on coefs

1/1.25

7. Plots:

label size (not too small) captions

placement pairs plots **NOT BLURRY**

0.25/1

8. Conclusions – make separate section

recap analysis * interpretation state main findings ↳ be specific

1/1

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

poor satisfactory good excellent

10. Other comments:

* cannot conclude causation, only causation

- no refs

- use your own words

2.5/4.25

good job!! ☺ Don't need to re-do +.25 NB

GLM: Name _____ AB 11.5/12 → 5.875/6 → 6/6

1. Formatting:

0.75/0.75

all margins 2.5cm

12 pt size

no raw R code or output

max **10** pages

informative title

name on all pages

all pages numbered

no blurry plots (**NOT png**)

2. Introduction/Background:

1/1

brief statement of scientific question

all variables defined

3. EDA:

2/2

univariate numerical

univariate graphical

bivariate numerical (cor)

bivariate graphical

4. Model fitting: ☺ very good

2/2

give mathematical definition of model

state how model fitted (ie, maximum likelihood)

CLEARLY describe how model selected

define all terms

5. Model assessment:

1.5/2

CLEARLY state model assumptions:

1. count outcome Poisson
2. independent obs
3. linear relation between log count and linear predictor
4. conditional mean = conditional variance

carry out assessment (numerical / graphics):
scatterplots (linearity assumption)

☺ model: $\log Y = X\beta$, est $\widehat{\log Y} = X\hat{\beta}$

not $\log \hat{Y}$

7.25/7.75

$\widehat{\log y}$ not $\log \hat{y}$

1/1 6. Write out final estimated model **mathematically**

hat on response variable (ok if coefs in table)

max **2 sig digits** on coefs

7. Plots:

1.25/1.25 label size (not too small)

captions

placement

NOT BLURRY

0.75/1 8. Conclusions

(*) interpretation

recap analysis

state main findings

1.25/1.25 9. Overall presentation (clarity of explanations, appropriate citations / references) :

poor satisfactory good excellent

10. Other comments:

(*) Careful: causation can be claimed for experimental study, otherwise only association

- when you have inline fractions, use \displaystyle so that the frac not too small

4.25/4.25

GLM: Name ED 9/12 → 4.5/6

1. Formatting:

0.5/0.75

all margins 2.5cm — Tables

12 pt size

no raw R code or output

max **10** pages

informative title

name on all pages

all pages numbered

no blurry plots (**NOT png**)

too many digits

2. Introduction/Background:

1/1

brief statement of scientific question

all variables defined

1.75/2

3. EDA: 5-number summary

univariate numerical

bivariate numerical (cor)

univariate graphical

bivariate graphical

4. Model fitting: I don't understand your model names

1.25/2

give mathematical definition of model

state how model fitted (ie, maximum likelihood)

CLEARLY describe how model selected - lowest AIC value in text does not agree with table

define all terms AIC integer df

5. Model assessment:

1/2

CLEARLY state model assumptions:

1. count outcome Poisson
2. independent obs
3. linear relation between log count and linear predictor
4. conditional mean = conditional variance

carry out assessment (numerical / graphics):
scatterplots (linearity assumption)

- Define all tests: null/alt/Test Stat/null dist/p
- Table 5 below description
- Don't need R fns

5.5/7.75

- why do you mention all coefs? It's confusing

0.5/1 - not red

6. Write out final estimated model **mathematically**
   - write in terms of $\ln y$
   hat on response variable max **2 sig digits** on coefs
   (ok if coefs in table)

1.25/1.25 - Don't shade p-values - hard to read

7. Plots:

label size (not too small)	captions
placement	**NOT BLURRY**

0.75/1 8. Conclusions
+EDA

recap analysis	(*) interpretation state main findings

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

poor satisfactory (good) excellent

10. Other comments:

- Don't need title page / table of contents

(*) cannot conclude causation ('gender does not impact'), only association

- use primary refs + data paper ref

3.5/4.25

GLM: Name CF 6.75/12 → 3.375/6

1. Formatting:

0.5/0.75

all margins 2.5cm — informative title

12 pt size — name on all pages

no raw R code or output — all pages numbered

max **10** pages — no blurry plots (**NOT png**)

0.5

2. Introduction/Background:

brief statement of scientific question

all variables defined

0.25/2

3. EDA: QQ?

univariate numerical — bivariate numerical (cor)

univariate graphical — bivariate graphical

all vars

don't need boxplot

all pairs

4. Model fitting:

somewhat incomplete

1.25/2

give mathematical definition of model

state how model fitted (ie, maximum likelihood)

CLEARLY describe how model selected

define all terms AIC

5. Model assessment:

1.5/2

CLEARLY state model assumptions:

1. count outcome Poisson
2. independent obs
3. linear relation between log count and linear predictor
4. conditional mean = conditional variance

?? linear predictor

carry out assessment (numerical / graphics):

scatterplots (linearity assumption)

SQUARE QQ

→ are you sure you did QQ for pois?

4/7.75 var = 4x mean!!

write equation numerically

0.75/1

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

hat on response variable (ok if coefs in table)

max **2 sig digits** on coefs

1/1.25

7. Plots:

label size (not too small)

captions

placement

NOT BLURRY

0/1

8. Conclusions not done

recap analysis

state main findings

1/1

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

poor satisfactory good excellent

10. Other comments:

- also cite data paper

- use your own words (background/data description)

2.75/4.25

GLM: Name AI 9.25/12 → 4.625/6

1. Formatting:

all margins 2.5cm	informative title
12 pt size captions?	name on all pages
no raw R code or output	all pages numbered
max **10** pages	no blurry plots (**NOT png**)

0.75/0.75

2. Introduction/Background:

brief statement of scientific question

all variables defined

1/1

3. EDA:

univariate numerical	bivariate numerical (cor)
univariate graphical	bivariate graphical scatterplot

no decimal for the whole numbers

1.25/2

4. Model fitting:

give mathematical definition of model – incorrectly specified*

state how model fitted (ie, maximum likelihood)

CLEARLY describe how model selected

define all terms

why not expect gender?

1.5/2

Could max L for AIC

Your model not min AIC if you include interactions

5. Model assessment:

Fig 5a QQ not hist

CLEARLY state model assumptions:

1. count outcome Poisson
2. independent obs
3. linear relation between log count and linear predictor
4. conditional mean = conditional variance

carry out assessment (numerical / graphics):
scatterplots (linearity assumption)

1.75/2

* model is $\log\lambda = \eta$, est is $\widehat{\log y} = \hat{\eta}$ (not $\log\bar{y}$)

6.25/7.75

write in terms of $\widehat{\ln y}$ (not $\hat{y}$)

0.75/1

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

hat on response variable (ok if coefs in table)

max **2 sig digits** on coefs

0.75/1.25

7. Plots:

label size (not too small)

captions too small!?

placement

NOT BLURRY

0.5/1

8. Conclusions

x-fold predicted increase

recap analysis

state main findings

↳ at the end somewhat vague and generic

1/1

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

poor satisfactory good excellent

10. Other comments:

- use primary resources (not Wikipedia)

- Figure (not fig.)

3/4.25

good job!! ☺ (*) Don't need to re-do +.25 train/test

GLM: Name TR 10.75/12 → 5.5/6 → 6/6

1. Formatting:

0.75/0.75

all margins 2.5cm	informative title
12 pt size	name on all pages
no raw R code or output	all pages numbered
max **10** pages	no blurry plots (**NOT png**)

2. Introduction/Background:

1/1

brief statement of scientific question

all variables defined

3. EDA:

2/2

univariate numerical	bivariate numerical (cor)
univariate graphical	bivariate graphical

4. Model fitting: (*) + clearly explain your 'evaluation' when you first mention it

1.75/2

give mathematical definition of model

state how model fitted (ie, maximum likelihood)

CLEARLY describe how model selected

define all terms

5. Model assessment: (9) needs ^ on all pars, + the table contains estimated par values

1.25/2 incomplete

CLEARLY state model assumptions:

1. count outcome Poisson
2. independent obs
3. linear relation between log count and linear predictor
4. conditional mean = conditional variance

carry out assessment (numerical / graphics):
scatterplots (linearity assumption)

(*) the model is $\log \lambda = X\beta$, est is $\widehat{\log Y} = X\hat{\beta}$ (not $\log \hat{Y}$)

6.75/7.75

1/1

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

hat on response variable (ok if coefs in table)

max **2 sig digits** on coefs

1.25/1.25

7. Plots:

label size (not too small)

captions

placement

NOT BLURRY

0.5/1

8. Conclusions

recap analysis

state main findings

1.25/1

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

poor satisfactory good excellent

10. Other comments:

- Don't need title page/contents/EPFL logo

- Poisson, not poisson

- use paragraphing, report hard to read

4/4.25

GLM: Name RP 8.5/12 → 4.25/6

1. Formatting: 0.75/0.75

all margins 2.5cm — informative title

12 pt size — name on all pages

no raw R code or output — all pages numbered

max **10** pages — no blurry plots (**NOT png**)

2. Introduction/Background: 0.75/1

brief statement of scientific question — say more

all variables defined (in EDA)

3. EDA: why fig1? Don't need Figure 3 1.5/2

univariate numerical — 5-number summary — bivariate numerical (cor)

univariate graphical — bivariate graphical

4. Model fitting: 1/2 ⊛

↳ Figure 2 label - incomplete

- why is tension L = intercept??

give mathematical definition of model

state how model fitted (ie, maximum likelihood)

CLEARLY describe how model selected

define all terms — quasi Poisson

5. Model assessment: Figure 5 should be square 1/2

CLEARLY state model assumptions:

1. count outcome Poisson
2. independent obs
3. linear relation between log count and linear predictor
4. conditional mean = conditional variance

carry out assessment (numerical / graphics):
scatterplots (linearity assumption)

incomplete

⊛ model is $\log Y = X\beta$, not is $\widehat{\log Y} = X\hat{\beta}$ not $\log \hat{Y}$

5/7.75

$\widehat{\log Y}$ not $\widehat{\log Y}$

1/1

6. Write out final estimated model **mathematically** ok given model misspecification

hat on response variable (ok if coefs in table)

max **2 sig digits** on coefs

4/1.25 7. Plots: make 'pretty' labels

label size (not too small)

captions

placement

NOT BLURRY

0.5/1 8. Conclusions

recap analysis

(*) interpretation

state main findings

very brief

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

poor satisfactory good excellent

10. Other comments:

(*) careful - can conclude causation from experimental study, otherwise only association

3.5/4.25

GLM: Name AS 8.75/12 → 4.375/6

1. Formatting:

0.75/0.75

- all margins 2.5cm
- 12 pt size
- **no raw R code or output** → R formulas in model tables
- max **10** pages
- informative title
- name on all pages
- all pages numbered
- no blurry plots (**NOT png**)

2. Introduction/Background:

1/1

- brief statement of scientific question
- all variables defined

3. EDA:

1/2 remove no pie charts !!!

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

4. Model fitting: Don't need R fns

1.25/2

- give mathematical definition of model – (1) not correct (*)
- state how model fitted (ie, maximum likelihood)
- CLEARLY describe how model selected bottom p. 4, refer specifically to Table 2
- define all terms AIC / IRLS

5. Model assessment: what hyp test for overdispersion?

Assess, not confirm

1.75/2

CLEARLY state model assumptions:

1. count outcome Poisson
2. independent obs
3. linear relation between log count and linear predictor
4. conditional mean = conditional variance – not proven

carry out assessment (numerical / graphics):
scatterplots (linearity assumption)

(*) model is $\log \lambda = X\beta$, est is $\widehat{\log Y} = X\hat{\beta}$ (not $\log \hat{Y}$)

5.75/7.75

$\ln(A)$ not $\ln(\hat{A})$

0.75/1 6. Write out final estimated model **mathematically**

inconsistent

hat on response variable (ok if coefs in table)

max **2 sig digits** on coefs

0.5/1.25 7. Plots: no pie charts!

label size (not too small)

captions

placement

NOT BLURRY

some plots too small

0.75/1 8. Conclusions

use paragraphs

not 'proven'

recap analysis

state main findings

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

poor satisfactory (good) excellent

10. Other comments:

- Header is unnecessary and distracting

3/4.25

good job!! ☺ Don't need to re-do

GLM: Name _____ HS 11.25/12 → 5.625/6 → 6/6

1. Formatting:

0.5/0.75

all margins 2.5cm

12 pt size

no raw R code or output

max **10** pages

informative title

name on all pages

all pages numbered

no blurry plots (**NOT png**)

– too many digits

2. Introduction/Background:

1/1

brief statement of scientific question

'vital' ??

all variables defined

3. EDA:

2/2

univariate numerical

univariate graphical

bivariate numerical (cor)

bivariate graphical

4. Model fitting:

1.75/2

give mathematical definition of model

state how model fitted (ie, maximum likelihood)

CLEARLY describe how model selected

define all terms AIC

5. Model assessment:

very good

2/2

CLEARLY state model assumptions:

1. count outcome Poisson
2. independent obs
3. linear relation between log count and linear predictor
4. conditional mean = conditional variance

carry out assessment (numerical / graphics):
scatterplots (linearity assumption)

☺ model is $\log Y = X\beta$, est is $\widehat{\log Y} = X\hat{\beta}$
not $\log \hat{Y}$

7.25/7.75

1/1 $\widehat{\log y}$ not $\log \hat{y}$

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

hat on response variable (ok if coefs in table)

max **2 sig digits** on coefs

1.25/1.25 7. Plots:

label size (not too small)

captions

placement

NOT BLURRY

0.5/1 8. Conclusions

+ EDA

recap analysis

(*) interpretation

state main findings

1.25/1.25 9. Overall presentation (clarity of explanations, appropriate citations / references) :

poor satisfactory good excellent

10. Other comments:

(*) careful - can conclude causation for experimental study, otherwise only association

4/4.25