

GLM: Name _____ **NB**

9.75/12 → 4.875/6

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)
<i>(too many digits)</i> |

2. Introduction/Background:

1/1

brief statement of scientific question

all variables defined

1.5/2

3. EDA:

univariate numerical

↳ 5-number summary

Fig 4 unnecessary explanation

univariate graphical

↳ don't need boxplots (Fig 2)

biivariate graphical

↳ all pairs

4. Model fitting:

↳ what vars are you using?

give mathematical definition of model

1.25/2

state how model fitted (ie, maximum likelihood)

- *Don't need numerical details*

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)

log-odds?

- ?? 2 numbers
are always
proportional

~~Sec. 2.1 goes in Model fitting~~
6/7.75

1/1 (Write coefs intine)

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

hat on response variable
(ok if coefs in table)

max **2 sig digits** on coefs

1.25 7. Plots: (Shapes) + incomplete

label size (not too small)

captions

placement

NOT BLURRY

0.5/1 8. Conclusions

(+ EDA
recap analysis)

somewhat vague at end
state main findings
'seems to be'?

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

poor

satisfactory

good

excellent

10. Other comments:

3.25/4.25

GLM: Name LB

5.5/12 → 2.75/6

1. Formatting:

0. 5 | 0.75
all margins 2.5cm
12 pt size

informative title

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:

use your own words

brief statement of scientific question

all variables defined

3. EDA:

univariate numerical

bivariate numerical (cor)

univariate graphical

bivariate graphical

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

5. Model assessment:

you assess not verify

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)

SQUARE QP

p. 46 top) very vague

3 | 2.75

0.25 | write numerical equation

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

hat on response variable
(ok if coeffs in table)

max 2 sig digits on coeffs

1/1.25 | 7. Plots:

label size (not too small)

captions

placement

pairs plots

NOT BLURRY

0.25 | 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:

① ~~*cannot conclude causation, only correlation~~

- no refs

- use your own words

2.5/4.25

good job!! *Don't need to re-do
GLM: Name AB $11.5/12 \rightarrow 5.875/6 \rightarrow 6/6$

1. Formatting:

- | | | |
|-----------------|--------------------------------|------------------------------------|
| 0.75 | all margins 2.5cm | informative title |
| 0.75 | 12 pt size | name on all pages |
| 0.75 | 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:

- ~~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: $\log \hat{Y} = X\beta$, est $\log \hat{Y} = X\hat{\beta}$
 $\log \hat{Y}$ not $\hat{\log Y}$

~~log y~~ not $\log y$

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

Y1

hat on response variable
(ok if coeffs in table)

max 2 sig digits on coeffs

7. Plots:

1.25 / 1.25

label size (not too small)
placement

captions

NOT BLURRY

0.75 / 1

8. Conclusions
recap analysis

④ interpretation
state main findings

1.25 / 1.25

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

poor satisfactory good

excellent

10. Other comments:

④ Careful: causation can be claimed for experimental study, otherwise only association

- when you have inline fractions, use displaystyle so that the frags 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
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

define all terms

AIC

- lowest AIC value in
text does not agree
with table

5. Model assessment:

integer df

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 // - not red

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

- write in terms of \hat{Y}

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

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

1.25 // 7. Plots:

label size (not too small)

captions

placement

NOT BLURRY

0.75 // 8. Conclusions

+ EDS

recap analysis

④ interpretation
state main findings

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

④ cannot conclude causation ('gender does not impact'), only association

- use primary refs + data paper ref

GLM: Name

CF

6.75 / 12 → 3.375 / 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 |

2. Introduction/Background:

brief statement of scientific question

all variables defined

0.25 / 2

3. EDA:

Q-Q plot

univariate numerical

univariate graphical

bivariate numerical (cor)

bivariate graphical

all vars
Don't need boxplot

all pairs

1.25 / 2

4. Model fitting:

somewhat incomplete

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 Q-Q

→ are you sure you did QQ for Pois?

4 / 2.75
Var = 4 × mean!!

write equation numerically

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

hat on response variable
(ok if coeffs in table)

max 2 sig digits on coeffs

7. Plots:

label size (not too small)

captions

placement

NOT BLURRY

8. Conclusions not done

recap analysis

state main findings

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 AJ

9.25/12 → 4.625/6

1. Formatting:

0.75/0.75
12 pt size
no raw R code or output

all margins 2.5cm

informative title

captions?

name on all pages

max **10** pages

all pages numbered

no blurry plots (NOT png)

1/1

2. Introduction/Background:

brief statement of scientific question

all variables defined

1.25/2

3. EDA:

univariate numerical

bivariate numerical (cor)

univariate graphical

bivariate graphical

scatterplot

no decimal for the whole numbers

1.5/2

4. Model fitting:

why not expect gender? - incorrectly specified
give mathematical definition of model

state how model fitted (ie, maximum likelihood)

CLEARLY describe how model selected

define all terms

Cooks max L for AIC

Your model not min AIC if you include interactions

1.75/2

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)

model is $\log \hat{y} = \eta$, est is $\log \hat{y} = \frac{1}{\eta}$ (not \log^{-1})

6.25/7.25

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

0.75 /

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

hat on response variable
(ok if coeffs in table)

max 2 sig digits on coeffs

0.75 /
1.25 /

7. Plots:

label size (not too small)
placement

captions
too small?

NOT BLURRY

0.5 /

8. Conclusions

recap analysis

x-fold predicted increase
state main findings
at the end somewhat vague
and generic

1 / /

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

poor

satisfactory

good

excellent

10. Other comments:

- use primary resources (not Wikipedia)
- Figures (not fig.)

good job!!  Don't need to re-do +.25 train/test
GLM: Name TK 10.75/12 → 5.5/6 → E/6

1. Formatting:

all margins 2.5cm	informative title
0.75/ 0.75 12 pt size	name on all pages
no raw R code or output	all pages numbered
max 10 pages	no blurry plots (NOT png)

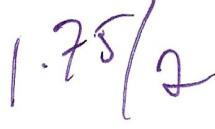
2. Introduction/Background:

	brief statement of scientific question
	all variables defined

3. EDA:

	univariate numerical	bivariate numerical (cor)
	univariate graphical	bivariate graphical

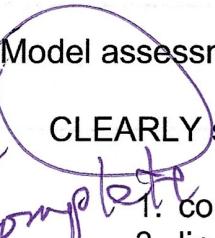
4. Model fitting:

	+ Clearly explain your 'evaluation' when you first mention it
	give mathematical definition of model
	state how model fitted (ie, maximum likelihood)

CLEARLY describe how model selected

define all terms

5. Model assessment:

	(a) needs ~ on all pairs + the table contains estimated peer values
	CLEARLY state model assumptions: <i>incomplete</i>

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

6.78 / 7.75

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

hat on response variable
(ok if coeffs in table)

max **2 sig digits** on coeffs

7. Plots:

label size (not too small)

captions

placement

NOT BLURRY

8. Conclusions

recap analysis

state main findings

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

9/4-25

GLM: Name _____

RP

8.5/12 → 4.25/6

1. Formatting:

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:

brief statement of scientific question

say more

all variables defined

(in EDA)

3. EDA: *why fig?* Don't need Figure 3

univariate numerical

bivariate numerical (cor)

univariate graphical

bivariate graphical

↳ Figure 2 label

4. Model fitting:

- *Why is tension L = intercept??* - incomplete
give mathematical definition of model

state how model fitted (ie, maximum likelihood)

CLEARLY describe how model selected

define all terms

quasi Poisson

Figure 5 should be square

5. Model assessment:

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 $\log Y = X\hat{\beta}$
not $\log \hat{Y}$

5/7.76

log Y not log \hat{Y}

6. Write out final estimated model mathematically

OK given model
misspecification

hat on response variable
(ok if coeffs in table)

max 2 sig digits on coeffs

7. Plots: make 'pretty' labels

label size (not too small)

captions

placement

NOT BLURRY

0.5
1
1

8. Conclusions

recap analysis

* Interpretation
state main findings
very brief

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

informative title

12 pt size

name on all pages

no raw R code or output

all pages numbered

max **10** pages

*R formulas in
model tables*

no blurry plots (NOT png)

2. Introduction/Background:

1/1

brief statement of scientific question

1/2

all variables defined

!!!

3. EDA:

~~remove~~ *no pie charts*

bivariate numerical (cor)

univariate graphical

bivariate graphical

1.25/2

4. Model fitting:

Don't need R fns

- (1) not correct

give mathematical definition of model

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?

Lassess, not confirm

1.25/2

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)

5.75/7.75

④ model is $\log \lambda = x\beta$, est is $\log \hat{Y} = X\hat{\beta}$ (not $\log Y$)

inconsistent

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

0.75

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

hat on response variable
(ok if coeffs in table)

max 2 sig digits on coeffs

0.5/
1.25

7. Plots: *no pie charts!*

label size (not too small)

captions

placement

NOT BLURRY

0.75 /

8. Conclusions

recap analysis

state main findings

*some plots too small
use paragraphs not 'proven'*

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!! \oplus Don't need to re-do

GLM: Name HS 11.25/12 \rightarrow 5.625/6 \rightarrow 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

'vital'??

2. Introduction/Background:

Y1

brief statement of scientific question

all variables defined

3. EDA:

2/2

univariate numerical

bivariate numerical (cor)

univariate graphical

bivariate graphical

4. Model fitting:

\ominus

give mathematical definition of model

1.75/
2

state how model fitted (ie, maximum likelihood)

CLEARLY describe how model selected

define all terms

AIC

5. Model assessment:

very good

CLEARLY state model assumptions:

2/2

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)

\oplus model is $\log Y = X\beta$, est is $\hat{\log Y} = \hat{X}\hat{\beta}$

$\hat{\log Y} = \hat{X}\hat{\beta}$
not $\log \hat{Y}$

2.25/
7.75

log Y not log \hat{Y}

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

hat on response variable
(ok if coeffs in table)

max 2 sig digits on coeffs

1.25
1.25

7. Plots:

label size (not too small)

captions

placement

NOT BLURRY

0.5/
1 EDA
recap analysis

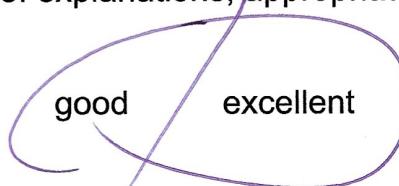
④ interpretation
state main findings

1.25
1.25

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

poor

satisfactory



10. Other comments:

④ careful - can conclude causation for
experimental study, otherwise only association

4/4.25