



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 7. Plots:

label size (not too small)

captions

1.25

placement

**NOT BLURRY**

8. Conclusions

(recap analysis)

state main findings

0.25 / 0.5

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

poor

satisfactory

good

excellent

0.25 / 1

10. Other comments:

- ref methods

- no RAW R!!

3.5

logistic: Name \_\_\_\_\_

Louise LARGEAU

7.75/12 → 3.875/6

1. Formatting:

all margins 2.5cm

informative title

12 pt size

name on all pages

no raw R code or output

all pages numbered

max 7 pages

no blurry plots (NOT png)

*write models mathematically not R formula*

0.25/0.5

2. Introduction/Background:

brief statement of scientific question

all variables defined

0.5/0.5

3. EDA:

univariate numerical

bivariate numerical (cor)

plot only 1 var

univariate graphical

bivariate graphical

4. Model fitting:

*- show all results in table: coef, SE, z, p*

give mathematical definition of model

state how model fitted (ie, maximum likelihood)

CLEARLY describe how model selected

define all terms

*f? Deviance z?*

1.25/2

5. Model assessment:

CLEARLY state model assumptions: *missing* + give PRIMARY reference

1. binary outcome

2. independent obs

3. linear relation between logit and linear predictor

4. no multicollinearity 5. no outliers (6. large sample size)

carry out assessment (numerical / graphics):

scatterplots of logit vs. predictors (linearity assumption) *missing*

did mention but did not define

DEFINE

Cook's distance / standardized residuals (outliers)

vif (to identify multicollinearity) *missing*

*Define*

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

7. Plots:

(label size) (not too small)  
placement

captions

**NOT BLURRY**

1.5 / 1.5

8. Conclusions

recap analysis

state main findings

0.75 / 0.5

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

poor

satisfactory

good

excellent

0.5 / 1

10. Other comments:

ref methods

---



---



---



---



---



---



---



---



---



---

3.75

logistic: Name

Claire Lathuy

7.75/12 → 3.875/6

1. Formatting:

- Don't need EPFL logo

all margins 2.5cm

informative title

12 pt size

name on all pages

no raw R code or output

all pages numbered

max 7 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:

give mathematical definition of model

state how model fitted (ie, maximum likelihood)

CLEARLY describe how model selected

define all terms

AIC

- how do you choose which var to add?

5. Model assessment:

ref missing, and should add more explanation

CLEARLY state model assumptions: + give PRIMARY reference

1. binary outcome

2. independent obs

3. linear relation between logit and linear predictor

4. no multicollinearity 5. no outliers (6. large sample size)

carry out assessment (numerical / graphics):

scatterplots of logit vs. predictors (linearity assumption)

DEFINE -> Cook's distance / standardized residuals (outliers)

vif (to identify multicollinearity)

report in text only saying "below 5", not enough

0.5/0.5

0.5/0.5

0.5/2

no boxplot?  
distribution of IV?

1.5/2

1.25/2

4.25

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

**hat** on response variable  
(ok if coefs in table)

max **2 sig digits** on coefs

7. Plots:

(label size) (not too small)

captions

placement

**NOT BLURRY**

8. Conclusions

recap analysis

(more concretely, starting in new paragraph)  
state main findings

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

poor

**satisfactory**

good

excellent

10. Other comments:

only give specific refs (not lecture notes)  
and cite in text

don't need to give results from each step,  
just explain clearly what is going on

logistic: Name

Tran Minh Son Le

9.25/12 → 4.625/6

1. Formatting:

all margins 2.5cm

informative title

12 pt size

name on all pages

no raw R code or output

all pages numbered

~~max 7 pages~~

no blurry plots (**NOT** png)

0.5/0.5

2. Introduction/Background:

brief statement of scientific question

all variables defined

0.5/0.5

3. EDA:

univariate numerical

bivariate numerical (cor)

univariate graphical

bivariate graphical  
I expect to see graphs showing corr, not response vs predictor

1.5/2

4. Model fitting:

give mathematical definition of model  
state how model fitted (ie, maximum likelihood)

not correctly specified

**CLEARLY** describe how model selected (backward)

define all terms

AIC

1.25/2

5. Model assessment:

**CLEARLY** state model assumptions: + give PRIMARY reference

1. binary outcome
2. independent obs
3. linear relation between logit and linear predictor
4. no multicollinearity
5. no outliers
6. large sample size

carry out assessment (numerical / graphics):

scatterplots of logit vs. predictors (linearity assumption)

**DEFINE ->** Cook's distance / standardized residuals (outliers)

vif (to identify multicollinearity)

a bit too short, can say more

5.25

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

7. Plots:

label size (not too small)

captions

placement

**NOT BLURRY**

1.5/1.5

8. Conclusions

recap analysis

state main findings

0.25/0.5

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

poor

satisfactory

good

excellent

0.5/1

10. Other comments:

-use primary references

---



---



---



---



---



---



---



---



---



---

y



Don't need to re-do

logistic: Name \_\_\_\_\_

George Lee

10.5/12

5.75/6 → 6/6

1. Formatting:

good job!

0.5/0.5

all margins 2.5cm

informative title

12 pt size

name on all pages

no raw R code or output

all pages numbered

~~max 7 pages~~

no blurry plots (NOT png)

2. Introduction/Background:

0.5/0.5

(ok)

looks ok, but could be better structured perhaps?

brief statement of scientific question

all variables defined

3. EDA:

2/2

univariate numerical

bivariate numerical (cor)

univariate graphical

bivariate graphical

density plots don't have nice shape

4. Model fitting:

2/2

good

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: + give PRIMARY reference

↻

1. binary outcome

2. independent obs

3. linear relation between logit and linear predictor

4. no multicollinearity 5. no outliers (6. large sample size)

carry out assessment (numerical / graphics):

also perfect fit, print (ok)

scatterplots of logit vs. predictors (linearity assumption)

DEFINE → Cook's distance / standardized residuals (outliers)

vif (to identify multicollinearity)

Define

6.5

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

hat on response variable  
(ok if coefs in table)

max **2 sig digits** on coefs

ok, but using conditional expectation

7. Plots:

(label size) (not too small)

placement

captions

**NOT BLURRY**

8. Conclusions

recap analysis

- incorrect interpretation to  
state main findings say 'affects',  
survival'

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

poor

satisfactory

good

excellent

10. Other comments:

- ref primary sources

good job!

4

logistic: Name Jiantian Lei 10.5/12 → 5.25/6

1. Formatting:

0.5/0.5

all margins 2.5cm

informative title

12 pt size

name on all pages

no raw R code or output

all pages numbered

~~max 7 pages~~

no blurry plots (**NOT** png)

0.5/0.5

2. Introduction/Background: (ok)

brief statement of scientific question **too short**

all variables defined

2/2

3. EDA:

univariate numerical

bivariate numerical (cor)

univariate graphical

bivariate graphical

1.5/2

4. Model fitting: include z-value in table

give mathematical definition of model

state how model fitted (ie, maximum likelihood)

CLEARLY describe how model selected

define all terms

AIC, stepwise, McFadden

5. Model assessment:

1.75/2

CLEARLY state model assumptions: + give PRIMARY reference

1. binary outcome

2. independent obs

3. linear relation between logit and linear predictor

4. no multicollinearity 5. no outliers (6. large sample size)

carry out assessment (numerical / graphics):

scatterplots of logit vs. predictors (linearity assumption)

DEFINE -> Cook's distance / standardized residuals (outliers)

missing defining vif (to identify multicollinearity)

Define - Cook not cook

6.25

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

7. Plots:

label size (not too small)

captions

placement

**NOT BLURRY**

1/1.5

8. Conclusions

recap analysis

state main findings

0.5/0.5

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

poor

satisfactory

good

excellent

1/1

10. Other comments:

- Fig 5 BEFORE conclusions

- good job

4.25

logistic: Name \_\_\_\_\_

Gabin Leroy

8.25/12 → 4.125/6

1. Formatting:

all margins 2.5cm

informative title

12 pt size

name on all pages

no raw R code or output

all pages numbered

max 7 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:

give mathematical definition of model

state how model fitted (ie, maximum likelihood)

**CLEARLY** describe how model selected

define all terms

5. Model assessment:

**CLEARLY** state model assumptions: + give **PRIMARY** reference

1. binary outcome

2. independent obs

3. linear relation between logit and linear predictor

4. no multicollinearity 5. no outliers (6. large sample size)

carry out assessment (numerical / graphics):

scatterplots of logit vs. predictors (linearity assumption)

**DEFINE** -> **Cook's distance** / standardized residuals (outliers)

vif (to identify multicollinearity)

0.5/0.5

0.5/0.5

2/2

1.5/2

0.75/2

5.25

Fig 1 shape not 'pretty'

OR Deviance incomplete

how choose vars to add

0.25 / 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

7. Plots:

label size (not too small)

captions

placement

**NOT BLURRY**

1 / 1.5

8. Conclusions

recap analysis

state main findings

0.25 / 0.5 / 0.5 / 1

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

poor

satisfactory

good

excellent

10. Other comments:

- ref data and methods

- model assessment incomplete

logistic: Name Hain Luud 7.5/12 → 3.25/6

1. Formatting:

0.5/0.5

all margins 2.5cm

informative title *haha!*

12 pt size

name on all pages

**no raw R code or output**

all pages numbered

max 7 pages

*Write models mathematically not R formulas*

no blurry plots (**NOT png**)

2. Introduction/Background:

0.5/0.5

brief statement of scientific question

all variables defined

3. EDA:

1.25/2

*dens. plots not 'pretty'*

only report 1 corr coef in text

univariate numerical

bivariate numerical (cor)

univariate graphical

bivariate graphical *why plotting logit vs predictor here?*

4. Model fitting:

1.25/2

give mathematical definition of model

state how model fitted (ie, maximum likelihood)

**CLEARLY** describe how model selected

define all terms *AIC, McFadden*

5. Model assessment:

*missing*

**CLEARLY** state model assumptions: + give **PRIMARY** reference

1. binary outcome

2. independent obs

3. linear relation between logit and linear predictor

4. no multicollinearity 5. no outliers (6. large sample size)

carry out assessment (numerical / graphics):

scatterplots of logit vs. predictors (linearity assumption)

**DEFINE ->** Cook's distance / standardized residuals (outliers)

vif (to identify multicollinearity)

no def or explanation

*incomplete*

4.25

0.5/1

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

**hat** on response variable  
(ok if coefs in table)

**max 2 sig digits** on coefs  
in a table, but not very detailed, and  
check for the column names

1/1

7. Plots:

label size (not too small)

captions

placement

**NOT BLURRY**

1/1.5

8. Conclusions

recap analysis

state main findings

0.25/0.5

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

poor

satisfactory

good

excellent

0.5/1

10. Other comments:

- cite primary refs

---



---



---



---



---



---



---



---



---



---

3.25



Don't need to re-do

logistic: Name \_\_\_\_\_

Antoine Maier

11/12 → 5.5/6

6/6

1. Formatting:

all margins 2.5cm

informative title

12 pt size

name on all pages

no raw R code or output

all pages numbered

~~max 7 pages~~

no blurry plots (**NOT png**)

0.5/0.5

2. Introduction/Background:

(ok)

brief statement of scientific question

Very few details in intro about the method that is used

all variables defined

0.5/0.5

3. EDA:

univariate numerical

bivariate numerical (cor)

univariate graphical

bivariate graphical

2/2

4. Model fitting:

give mathematical definition of model

state how model fitted (ie, maximum likelihood)

CLEARLY describe how model selected

define all terms

AIC

2/2

5. Model assessment:

CLEARLY state model assumptions: + give PRIMARY reference

1. binary outcome

2. independent obs

3. linear relation between logit and linear predictor

4. no multicollinearity 5. no outliers (6. large sample size)

1.5/2

carry out assessment (numerical / graphics):

scatterplots of logit vs. predictors (linearity assumption)

DEFINE -> Cook's distance / standardized residuals (outliers) No plot for resid.

only one numerical cov. vif (to identify multicollinearity)

6.5

1/1

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

~~why hat on eta??~~

**hat** on response variable  
(ok if coefs in table)

max **2 sig digits** on coefs

1/1

7. Plots:

label size (not too small)

captions

placement

**NOT BLURRY**

1/1.5

8. Conclusions

*(expand  
recap analysis)*

*(start in new paragraph  
state main findings)*

0.5/0.5

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

poor

satisfactory

good

excellent

1/1

10. Other comments:

*good job!*

---



---



---



---



---



---



---



---



---



---

4.5

Don't need to re-do

logistic: Name \_\_\_\_\_

Malvar Fonseca Vilela

10/12 → 5/6

→ 6/6  
good job!

1. Formatting:

0.5/0.5

all margins 2.5cm

informative title

12 pt size

name on all pages

no raw R code or output

all pages numbered

max 7 pages

no blurry plots (NOT png)

2. Introduction/Background:

0.5/0.5

brief statement of scientific question

all variables defined

3. EDA:

1.5/2

univariate numerical

bivariate numerical (cor)

univariate graphical

bivariate graphical

4. Model fitting:

1.5/2

give mathematical definition of model

state how model fitted (ie, maximum likelihood)

CLEARLY describe how model selected

define all terms

5. Model assessment:

1.75/2

CLEARLY state model assumptions: + give PRIMARY reference

1. binary outcome

2. independent obs

3. linear relation between logit and linear predictor

4. no multicollinearity 5. no outliers (6. large sample size)

carry out assessment (numerical / graphics):

scatterplots of logit vs. predictors (linearity assumption)

DEFINE -> Cook's distance / standardized residuals (outliers)

vif (to identify multicollinearity)

5.75

logit

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

7. Plots:

(label size (not too small)

captions

placement

**NOT BLURRY**

1.5 / 1.5

8. Conclusions

recap analysis

state main findings

0.25 / 0.5

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

poor

satisfactory

good

excellent

0.75 / 1

10. Other comments:

- ref data + methods

- good job, I like your node-age interpretation

4.25

logistic: Name \_\_\_\_\_

Blanche MARION

8/12 → 4/6

1. Formatting:

all margins 2.5cm

informative title

0.5/0.5

12 pt size

name on all pages

no raw R code or output

all pages numbered

max 7 pages

no blurry plots (NOT png)

- too many digits  
(not detailed enough)

2. Introduction/Background:

0.5/0.5

brief statement of scientific question

all variables defined

3. EDA:

1/2

univariate numerical

bivariate numerical (cor)

univariate graphical

bivariate graphical

4. Model fitting:

1.5/2

give mathematical definition of model

state how model fitted (ie, maximum likelihood)

- MLE not 'algorithm' in the traditional sense

CLEARLY describe how model selected

define all terms

Describe what you are doing  
odds ratio

5. Model assessment:

1.5/2

CLEARLY state model assumptions: + give PRIMARY reference

EXPLAIN how the graphs help to assess assumptions

1. binary outcome

2. independent obs

3. linear relation between logit and linear predictor

4. no multicollinearity 5. no outliers (6. large sample size)

carry out assessment (numerical / graphics):

scatterplots of logit vs. predictors (linearity assumption)

DEFINE -> Cook's distance / standardized residuals (outliers)

vif (to identify multicollinearity)

5

0.5

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

**hat** on response variable  
(ok if coefs in table)

max **2 sig digits** on coefs

1/1

7. Plots:

(label size (not too small)  
placement

captions

**NOT BLURRY**

0.75 / 1.5

8. Conclusions

recap analysis

state main findings

0.25 / 0.5 / 0.5 / 1

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

poor

satisfactory

good

excellent

10. Other comments:

- ref data and methods at end and cite

in text

---

---

---

---

---

---

---

---

---

---

logistic: Name \_\_\_\_\_

Marko Mekjavic

5.25 / 12 → 2.625 / 6

1. Formatting:

0.5 / 0.5

all margins 2.5cm

12 pt size

**no raw R code or output**

~~max 7~~ pages

(informative title)

name on all pages

all pages numbered

no blurry plots (**NOT png**)

2. Introduction/Background:

0.25 / 0.5

brief statement of scientific question

all variables defined in text, hard to read and no details

3. EDA:

not done?

why analyzing response var?

univariate numerical

bivariate numerical (cor)

univariate graphical

bivariate graphical title missing

4. Model fitting:

define logit

give mathematical definition of model

state how model fitted (ie, maximum likelihood)

**CLEARLY** describe how model selected

define all terms

5. Model assessment:

all missing

**CLEARLY** state model assumptions: + give PRIMARY reference

1. binary outcome

2. independent obs

3. linear relation between logit and linear predictor

4. no multicollinearity 5. no outliers (6. large sample size)

carry out assessment (numerical / graphics):

scatterplots of logit vs. predictors (linearity assumption)

**DEFINE** -> Cook's distance / standardized residuals (outliers)

vif (to identify multicollinearity)

2.25  
\* you can not leave out colors - if you use one you need them all

0.5/1

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

**hat** on response variable  
(ok if coefs in table)

max **2 sig digits** on coefs

1/1

7. Plots:

label size (not too small)

captions

placement

**NOT BLURRY**

0.75/1.5

8. Conclusions

recap analysis

(expand)  
state main findings

0.25/1

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

poor

satisfactory

good

excellent

0.5/1  
0.5/1

10. Other comments:

- logistic not linear regression

- refs for methods



logistic: Name \_\_\_\_\_

Elisa Michalski

6.75/12 → 3.375/6

1. Formatting:

all margins 2.5cm

informative title

12 pt size

name on all pages

no raw R code or output

all pages numbered

max 7 pages

no blurry plots (NOT png)

write models mathematically  
not R formula

2. Introduction/Background:

brief statement of scientific question

too short, few details

all variables defined

(ok)

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

5. Model assessment:

CLEARLY state model assumptions: + give PRIMARY reference

1. binary outcome
2. independent obs
3. linear relation between logit and linear predictor
4. no multicollinearity
5. no outliers
6. large sample size

carry out assessment (numerical / graphics):

scatterplots of logit vs. predictors (linearity assumption)

DEFINE -> Cook's distance / standardized residuals (outliers)

vif (to identify multicollinearity)

3.75

0.5 / 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

7. Plots:

label size (not too small)

captions

placement

**NOT BLURRY**

0.75 / 1.0

8. Conclusions

recap analysis

*expand*  
state main findings

0.25 / 0.5

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

poor

satisfactory

good

excellent

0.5 / 1

10. Other comments:

*- ref data + methods*

---



---



---



---



---



---



---



---



---



---

Don't need to re-do

logistic: Name \_\_\_\_\_

Gioele Monopoli

10.25/12 → 5.125/6

6/6  
good job!

1. Formatting:

0.5/0.5

all margins 2.5cm

12 pt size

no raw R code or output

max 7 pages

informative title

name on all pages

all pages numbered

no blurry plots (NOT png)

2. Introduction/Background:

0.5/0.5

brief statement of scientific question

all variables defined

3. EDA:

2/2

univariate numerical

univariate graphical  
some in appendix

bivariate numerical (cor)

OK bivariate graphical

← not sure I see this

4. Model fitting:

1.5/2

give mathematical definition of model

state how model fitted (ie, maximum likelihood)

CLEARLY describe how model selected

define all terms

AIC

- you use p to mean 2 different things  
Explain the procedure, not the R fn.

5. Model assessment:

1.75/2

CLEARLY state model assumptions:

+ give PRIMARY reference

1. binary outcome

2. independent obs

3. linear relation between logit and linear predictor

4. no multicollinearity 5. no outliers (6. large sample size)

carry out assessment (numerical / graphics):

scatterplots of logit vs. predictors (linearity assumption)

DEFINE → Cook's distance / standardized residuals (outliers)

vif (to identify multicollinearity)

Define

6.25

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

7. Plots:

label size (not too small)

captions

placement

**NOT BLURRY**

1.25 / 1.5

8. Conclusions

(*expand*)  
recap analysis

state main findings

0.25 / 0.5

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

poor

satisfactory

good

excellent

0.25 / 1

10. Other comments:

- ref data and methods

- good job!

4

logistic: Name \_\_\_\_\_

Meriadec Morrier

4.75/12 → 2.375/6

2.375/6

### 1. Formatting:

0.25/0.5

all margins 2.5cm

informative title

12 pt size

name on all pages

no raw R code or output

all pages numbered

max 7 pages

write models mathematically  
not R formula

no blurry plots (**NOT** png)

### 2. Introduction/Background:

0.5/0.5

brief statement of scientific question

all variables defined

### 3. EDA:

on predictors

univariate numerical

bivariate numerical (cor)

analyzed response var

univariate graphical

bivariate graphical

### 4. Model fitting:

- too many digits

0.75/2

give mathematical definition of model

state how model fitted (ie, maximum likelihood)

CLEARLY describe how model selected

define all terms

LRT, Deviance, AIC

### 5. Model assessment:

all missing

CLEARLY state model assumptions: + give PRIMARY reference

1. binary outcome

2. independent obs

3. linear relation between logit and linear predictor

4. no multicollinearity 5. no outliers (6. large sample size)

carry out assessment (numerical / graphics):

scatterplots of logit vs. predictors (linearity assumption)

**DEFINE** -> Cook's distance / standardized residuals (outliers)

vif (to identify multicollinearity)

2

0.5 / 1

too many digits

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

hat on response variable  
(ok if coefs in table)

max 2 sig digits on coefs

in a table, but very messy

1 / 1

7. Plots:

label size (not too small)

captions

placement

**NOT BLURRY**

0.75 / 1.5

8. Conclusions

(Don't bold)

recap analysis

(expand state main findings)

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

poor

satisfactory

good

excellent

0.5 / 1

10. Other comments:

- ref data + methods

- very incomplete

2.75

logistic: Name \_\_\_\_\_

Ni Erchang

7.75 / 12 → 3.875 / 6

1. Formatting:

all margins 2.5cm

informative title

12 pt size

name on all pages

no raw R code or output

all pages numbered

~~max 7 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 missing

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

AIC, ROC

5. Model assessment:

CLEARLY state model assumptions: + give PRIMARY reference

1. binary outcome

2. independent obs

3. linear relation between logit and linear predictor

4. no multicollinearity 5. no outliers (6. large sample size)

carry out assessment (numerical / graphics):

scatterplots of logit vs. predictors (linearity assumption)

DEFINE -> Cook's distance / standardized residuals (outliers)

vif (to identify multicollinearity)

- too many digits

3.75

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

7. Plots:

label size (not too small)

captions

placement

**NOT BLURRY**

1.5 / 1.5

8. Conclusions

recap analysis

state main findings

0.25 / 0.5

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

poor

satisfactory

good

excellent

0.5 / 1

10. Other comments:

- use primary references

- write models mathematically, not using R formulas

- too many digits



logistic: Name

Andreea Nica

8.75/12 → 4.375/6

### 1. Formatting:

all margins 2.5cm

informative title

12 pt size

name on all pages

no raw R code or output

all pages numbered

max 7 pages

no blurry plots (**NOT** png)

### 2. Introduction/Background:

brief statement of scientific question

too short

all variables defined

put in text, hard to read, did not specify type, coding, etc

### 3. EDA:

univariate numerical

bivariate numerical (cor)

univariate graphical

better separate

different graphs

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

### 5. Model assessment:

CLEARLY state model assumptions: + give PRIMARY reference

1. binary outcome

2. independent obs

3. linear relation between logit and linear predictor

4. no multicollinearity 5. no outliers (6. large sample size)

carry out assessment (numerical / graphics):

scatterplots of logit vs. predictors (linearity assumption)

DEFINE -> Cook's distance / standardized residuals (outliers)

vif (to identify multicollinearity)

5.5

0.5/1

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

**hat** on response variable  
(ok if coefs in table)

max **2 sig digits** on coefs

in a table

7. Plots:

1/1

label size (not too small)

captions

placement

**NOT BLURRY**

1/1.5

8. Conclusions

(*expand*)  
recap analysis

(*expand*)  
state main findings

0.25/0.5

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

poor

satisfactory

good

excellent

0.5/1

10. Other comments:

- refs (data + methods)

---



---



---



---



---



---



---



---



---



---

3.25

logistic: Name \_\_\_\_\_

Ioan Nitu

9.5/12 → 4.75/6

1. Formatting:

all margins 2.5cm

informative title

0.5/0.5 → 12 pt size

name on all pages

no raw R code or output

all pages numbered

max 7 pages

no blurry plots (NOT png)

2. Introduction/Background: (OK)

0.5/0.5

brief statement of scientific question

all variables defined

3. EDA:

Don't need

1/2

univariate numerical

bivariate numerical (cor)

Piecharts !!!

univariate graphical

bivariate graphical

histograms instead of boxplots

4. Model fitting:

1.75/2

give mathematical definition of model - no R formula

state how model fitted (ie, maximum likelihood)

CLEARLY describe how model selected

define all terms

Logit ← widehat

5. Model assessment:

1.5/2

CLEARLY state model assumptions: + give PRIMARY reference

1. binary outcome

2. independent obs

3. linear relation between logit and linear predictor

4. no multicollinearity 5. no outliers (6. large sample size)

carry out assessment (numerical / graphics):

strange? →

scatterplots of logit vs. predictors (linearity assumption)

DEFINE →

Cook's distance / standardized residuals (outliers)

vif (to identify multicollinearity)

5.25

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

1/1

**hat** on response variable  
(ok if coefs in table)

max **2 sig digits** on coefs

7. Plots:

1/1

label size (not too small)

captions

placement

**NOT BLURRY**

8. Conclusions

1.25/1.5

recap analysis

start new paragraph  
state main findings

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

0.25/0.5

poor

satisfactory

good

excellent

10. Other comments:

0.75

NO PIE CHARTS !!!

primary refs

4.25

logistic: Name \_\_\_\_\_

Aleksandra Novikova

7/12 → 3.5/6

1. Formatting:

all margins 2.5cm

informative title

12 pt size

name on all pages

**no raw R code or output**

all pages numbered

max 7 pages

no blurry plots (**NOT png**)

open quote " not "

2. Introduction/Background:

brief statement of scientific question **too short**

(OK)

all variables defined

3. EDA:

univariate numerical

bivariate numerical (cor)

univariate graphical

bivariate graphical

4. Model fitting:

give mathematical definition of model

- make this more clear

state how model fitted (ie, maximum likelihood)

(CLEARLY) describe how model selected

define all terms

AIC (mathematically)

5. Model assessment:

CLEARLY state model assumptions: **missing** + give PRIMARY reference

1. binary outcome

2. independent obs

3. linear relation between logit and linear predictor

4. no multicollinearity 5. no outliers (6. large sample size)

carry out assessment (numerical / graphics):

scatterplots of logit vs. predictors (linearity assumption)

**DEFINE ->** Cook's distance / standardized residuals (outliers)

vif (to identify multicollinearity)

no explanation of Cook's D

4.25

write as equation

0.5/1

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

**hat** on response variable

max **2 sig digits** on coeffs

(ok if coeffs in table) **in a table, very short though**

1/1

7. Plots:

(label size) (not too small)

captions

placement

**NOT BLURRY**

0.5/1.5

8. Conclusions

~~not done~~

(recap analysis)

(expand)

state main findings

0.25/0.5

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

poor

satisfactory

good

excellent

0.5/1

10. Other comments:

- use a spell checker

- primary refs for methods

- incomplete model assessment

2.75

Don't need to re-do

logistic: Name

Anna Paulish

10/12

→ 5/6 → 6/6

1. Formatting:

0.5 / 0.5

all margins 2.5cm

12 pt size

no raw R code or output

max 7 pages

informative title

name on all pages

all pages numbered

no blurry plots (NOT png)

good job!

2. Introduction/Background:

(OK)

0.5 / 0.5

brief statement of scientific question

all variables defined

Ok, a bit more details would help to read

3. EDA:

2 / 2

univariate numerical

bivariate numerical (cor)

univariate graphical

bivariate graphical (square)

4. Model fitting:

1.5 / 2

give mathematical definition of model

state how model fitted (ie, maximum likelihood)

**CLEARLY** describe how model selected

define all terms

5. Model assessment:

**CLEARLY** state model assumptions: + give PRIMARY reference

1.75 / 2

- 1. binary outcome
- 2. independent obs
- 3. linear relation between logit and linear predictor
- 4. no multicollinearity
- 5. no outliers
- 6. large sample size

carry out assessment (numerical / graphics):

scatterplots of logit vs. predictors (linearity assumption)

**DEFINE ->** Cook's distance / standardized residuals (outliers)

vif (to identify multicollinearity)

6.25

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

Not clearly written, uses conditional prob.

4 / 1

7. Plots:

label size (not too small)

captions

placement

**NOT BLURRY**

1 / 1.5

8. Conclusions

(expand)  
recap analysis

state main findings

0.25 / 0.5

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

poor

satisfactory

good

excellent

0.75 / 1

10. Other comments:

good job!

---



---



---



---



---



---



---



---



---



---

3.75



logistic: Name \_\_\_\_\_

Colin Pelletier

8/12 → 4/6

1. Formatting:

all margins 2.5cm

informative title

12 pt size

name on all pages

no raw R code or output

all pages numbered

max 7 pages

no blurry plots (NOT png)

2. Introduction/Background:

brief statement of scientific question

A bit short, not many details?

all variables defined

3. EDA:

not for all pairs

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

5. Model assessment:

CLEARLY state model assumptions: + give PRIMARY reference

1. binary outcome

2. independent obs

3. linear relation between logit and linear predictor

4. no multicollinearity 5. no outliers (6. large sample size)

carry out assessment (numerical / graphics):

scatterplots of logit vs. predictors (linearity assumption)

DEFINE -> Cook's distance / standardized residuals (outliers)

vif (to identify multicollinearity)

Define

4.75

0.75  
Y, and Y^,  
notation clash

Clarify

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

**hat** on response variable  
(ok if coefs in table)

max **2 sig digits** on coefs

1/1

7. Plots:

label size (not too small)

captions

placement

**NOT BLURRY**

0.75 / 1.5

8. Conclusions

recap analysis

expand  
state main findings

0.25 / 0.5

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

poor

satisfactory / good

excellent

0.5 / 1

10. Other comments:

---



---



---



---



---



---



---



---



---



---

3.25

Don't need to re-do

logistic: Name Lazar Radojević

10/12 → ~~5/6~~ → 6/6

good job!

1. Formatting:

all margins 2.5cm

informative title

0.5/0.5

12 pt size

name on all pages

no raw R code or output

all pages numbered

max 7 pages

no blurry plots (NOT png)

2. Introduction/Background:

0.25/0.5

brief statement of scientific question

all variables defined *too brief, need to explain the vars*

3. EDA:

2/2

univariate numerical

bivariate numerical (cor)

univariate graphical *missing*

bivariate graphical

4. Model fitting:

1.5/2

give mathematical definition of model

state how model fitted (ie, maximum likelihood)

(CLEARLY) describe how model selected

define all terms *AIC*

*- Table 3 hard to read (square)  
- These results should come AFTER model description*

5. Model assessment:

1.75/2

CLEARLY state model assumptions: *+ give PRIMARY reference*

- 1. binary outcome
- 2. independent obs
- 3. linear relation between logit and linear predictor
- 4. no multicollinearity
- 5. no outliers
- 6. large sample size

carry out assessment (numerical / graphics):

scatterplots of logit vs. predictors (linearity assumption)

DEFINE -> *Cook's distance* / standardized residuals (outliers)

vif (to identify multicollinearity)

*Define*

0.75 / logit

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

**hat** on response variable  
(ok if coefs in table)

max **2 sig digits** on coefs

1 / 1 / 7. Plots:

(label size (not too small)

captions

placement

**NOT BLURRY**

1.25 / 8. Conclusions

(expand  
recap analysis)

state main findings

0.25 / 0.5 / 9. Overall presentation (clarity of explanations, appropriate citations / references) :

poor      satisfactory      good      excellent

0.75 / 10. Other comments:

- use primary references (data + methods)

---

---

---

---

---

---

---

---

---

---

logistic: Name

Madeleine Robert

9.5/12 → 4.75/6

1. Formatting:

all margins 2.5cm

informative title

0.5/0.5

12 pt size

name on all pages

**no raw R code or output**

all pages numbered

~~max 7 pages~~

no blurry plots (**NOT png**)

0.5/0.5

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

4. Model fitting:

no ^ in mathematical description (!)

give mathematical definition of model

state how model fitted (ie, maximum likelihood)

**CLEARLY** describe how model selected

define all terms

ROC

5. Model assessment:

**CLEARLY** state model assumptions: + give PRIMARY reference

1. binary outcome

2. independent obs

3. linear relation between logit and linear predictor

4. no multicollinearity 5. no outliers (6. large sample size)

carry out assessment (numerical / graphics):

scatterplots of logit vs. predictors (linearity assumption)

**DEFINE** -> Cook's distance / standardized residuals (outliers)

vif (to identify multicollinearity)

should give more details

+ Define

^ on coefs

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

7. Plots:

label size (not too small)

captions

placement

**NOT BLURRY**



1/1.5

8. Conclusions

(*expand*  
recap analysis)

state main findings

0.25/0.5

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

poor

satisfactory

**good**

excellent



0.75/1

10. Other comments:

- cite primary refs for methods

---

---

---

---

---

---

---

---

---

---

3.5

logistic: Name \_\_\_\_\_

David Rochinha Chaves

6.75/12 → 3.375/6

1. Formatting:

all margins 2.5cm

informative title

0.5/0.5

12 pt size

name on all pages

no raw R code or output

all pages numbered

max 7 pages

no blurry plots (**NOT png**)

0.25/0.5

2. Introduction/Background:

brief statement of scientific question

Mentions main study, but doesn't say much what the project is about, in the sense that it follows the same approach as in the main paper

all variables defined

3. EDA:

modeling not in EDA

univariate numerical

bivariate numerical (cor)

univariate graphical

bivariate graphical

- too many digits

4. Model fitting:

give mathematical definition of model

THEN DO fitting

state how model fitted (ie, maximum likelihood)

CLEARLY describe how model selected

define all terms

logit, Deviance

5. Model assessment:

incomplete

CLEARLY state model assumptions: + give PRIMARY reference

1. binary outcome

2. independent obs

3. linear relation between logit and linear predictor

4. no multicollinearity 5. no outliers (6. large sample size)

carry out assessment (numerical / graphics):

scatterplots of logit vs. predictors (linearity assumption)

DEFINE -> Cook's distance / standardized residuals (outliers)

vif (to identify multicollinearity)

0.5/1

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

**hat** on response variable  
(ok if coefs in table)

max **2 sig digits** on coefs

1/1

7. Plots:

label size (not too small)

captions

placement

**NOT BLURRY**

0.5/1

8. Conclusions

*incomplete*

*expand*  
(state main findings)

recap analysis

0.25/0.5

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

poor

satisfactory

good

excellent

0.5/1

10. Other comments:

*- refs (data + methods)*

---



---



---



---



---



---



---



---



---



---

2.75



logistic: Name Romanenko 7/12 → 3.5/6

1. Formatting: interline too small, hard to read

- all margins 2.5cm
- informative title
- 12 pt size
- name on all pages
- 0.5/0.5 no raw R code or output
- all pages numbered
- max 7 pages
- no blurry plots (**NOT** png)

0.5/0.5 2. Introduction/Background:  
brief statement of scientific question

all variables defined

0.5/ 3. EDA: Don't need Shapiro-Wilk

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

1/2 4. Model fitting: mostly missing scatter plots  
give mathematical definition of model - Define logit

state how model fitted (ie, maximum likelihood)

CLEARLY describe how model selected Describe method not just for used  
define all terms AIC, logit

5. Model assessment: much is missing here

CLEARLY state model assumptions: + give PRIMARY reference

1/2

- 1. binary outcome
- 2. independent obs
- 3. linear relation between logit and linear predictor
- 4. no multicollinearity
- 5. no outliers
- 6. large sample size

carry out assessment (numerical / graphics):

scatterplots of logit vs. predictors (linearity assumption)

DEFINE -> Cook's distance / standardized residuals (outliers)  
vif (to identify multicollinearity)

Don't understand your interpretation of fig 3

3.5 Define

0.75

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

7. Plots:

label size (not too small)

captions

placement

**NOT BLURRY**

1.25 / 1.5

8. Conclusions

(*expand*)  
recap analysis

state main findings

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

0.25

0.5

poor

**satisfactory**

good

excellent

0.5 / 1

10. Other comments:

- use spell-checker

- ref data + methods

- somewhat incomplete

3.5

logistic: Name \_\_\_\_\_

Fabian Roulin

6.5/12 → 3.25/6

1. Formatting:

0.5/0.5

all margins 2.5cm

12 pt size

**no raw R code or output**

max 7 pages

open quote " not ))

informative title

name on all pages

all pages numbered

no blurry plots (**NOT png**)

2. Introduction/Background:

0.5/1

brief statement of scientific question

all variables defined

smth is off, it says linear regression

logistique

3. EDA:

1.5/2

univariate numerical

bivariate numerical (cor)

univariate graphical

bivariate graphical

4. Model fitting:

incomplete

0.75/2

give mathematical definition of model

state how model fitted (ie, maximum likelihood)

**CLEARLY** describe how model selected

define all terms

5. Model assessment:

0/2

not done

again in reference to linear regression

**CLEARLY** state model assumptions: + give **PRIMARY** reference

1. binary outcome
2. independent obs
3. linear relation between logit and linear predictor
4. no multicollinearity
5. no outliers
6. large sample size

carry out assessment (numerical / graphics):

scatterplots of logit vs. predictors (linearity assumption)

**DEFINE** -> Cook's distance / standardized residuals (outliers)

vif (to identify multicollinearity)

3.25

write as equation

0.5/1

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

**hat** on response variable  
(ok if coefs in table)

**max 2 sig digits** on coefs

0.5/1

7. Plots: Each graph should have a number and

(label size) (not too small)

captions

caption

placement

**NOT BLURRY**

1.5/1.5

8. Conclusions

recap analysis

state main findings

0.25/0.0

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

poor

**satisfactory**

good

excellent

0.5/1

10. Other comments:

- ref methods

- very incomplete

3.25

logistic: Name \_\_\_\_\_

Adrien Sizaret

6.75/12 →

3.375/6

1. Formatting:

all margins 2.5cm

informative title

0.5/0.5

12 pt size

name on all pages

no raw R code or output

all pages numbered

max 7 pages

no blurry plots (NOT png)

2. Introduction/Background:

0.25/0.5

brief statement of scientific question

Not many details

all variables defined

3. EDA:

Don't need QQ

0.25/2

univariate numerical

bivariate numerical (cor)

univariate graphical

bivariate graphical

cor matrix, not test

not sure if done properly

4. Model fitting:

- choose which variables (square not data) to include

0.75/2

incomplete

give mathematical definition of model not done

state how model fitted (ie, maximum likelihood)

CLEARLY describe how model selected

- alternative incorrectly specified

define all terms

- bottom p. 4 hard to read, write in terms of logit

5. Model assessment:

CLEARLY state model assumptions: + give PRIMARY reference

1/2

1. binary outcome

2. independent obs

3. linear relation between logit and linear predictor

4. no multicollinearity 5. no outliers (6. large sample size)

carry out assessment (numerical / graphics):

scatterplots of logit vs. predictors (linearity assumption)

given in a table??

DEFINE -> Cook's distance / standardized residuals (outliers)

vif (to identify multicollinearity)

Plots, not list of numbers

3.25

Explanations incomplete

0.5/1

$\logit(\pi)$

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

**hat** on response variable  
(ok if coefs in table)

**max 2 sig digits** on coefs

1/1

7. Plots:

(label size) (not too small)

captions

placement

**NOT BLURRY**

1.25/

8. Conclusions

1.5 (expand)  
recap analysis

state main findings

0.25/0.5

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

poor

satisfactory

good

excellent

0.5/1

10. Other comments:

- refs (data + methods)

- explanations incomplete

3.5

Dont need to re-do

logistic: Name \_\_\_\_\_

Haoxin Sun

10.25/12 → 75.125/6

→ 6/6  
good job!

1. Formatting:

0.5/0.5

all margins 2.5cm

informative title

12 pt size

name on all pages

**no raw R code or output**

all pages numbered

max 7 pages

no blurry plots (**NOT png**)

2. Introduction/**Background**:

(ok)

0.5/0.5

brief statement of scientific question

few details only

all variables defined

3. EDA:

1.5/2

**univariate numerical**

bivariate numerical (cor)

univariate graphical

bivariate graphical (square)

↳ BEFORE model fitting

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

5. Model assessment:

1.75/2

CLEARLY state model assumptions: + give **PRIMARY** reference

1. binary outcome

2. independent obs

3. linear relation between logit and linear predictor

4. no multicollinearity 5. no outliers (6. large sample size)

carry out assessment (numerical / graphics): see only wrt one covariate scatterplots of logit vs. predictors (linearity assumption)

**DEFINE** -> Cook's distance / standardized residuals (outliers)

vif (to identify multicollinearity)

1/1

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

**hat** on response variable  
(ok if coefs in table)

max **2 sig digits** on coefs

1/1

7. Plots:

label size (not too small)

captions

placement

**NOT BLURRY**

1.25 / 1.5

8. Conclusions

(*expand*  
*recap analysis*)

state main findings

0.5 / 0.5

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

poor

satisfactory

good

**excellent**

1/1

10. Other comments:

*good job!*

---



---



---



---



---



---



---



---



---



---

1.25



logistic: Name \_\_\_\_\_

Maria Tager

8.5/12 → 4.25/6

### 1. Formatting:

all margins 2.5cm

informative title

12 pt size

name on all pages

no raw R code or output

all pages numbered

max 7 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:

give mathematical definition of model

state how model fitted (ie, maximum likelihood)

CLEARLY describe how model selected

define all terms

### 5. Model assessment:

CLEARLY state model assumptions: + give PRIMARY reference

1. binary outcome

2. independent obs

3. linear relation between logit and linear predictor

4. no multicollinearity 5. no outliers (6. large sample size)

carry out assessment (numerical / graphics):

scatterplots of logit vs. predictors (linearity assumption)

DEFINE -> Cook's distance / standardized residuals (outliers)

vif (to identify multicollinearity)

Too many digits

5.29

write an equation

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

7. Plots:

(label size) (not too small)

captions

placement

**NOT BLURRY**

0.75/1

8. Conclusions

recap analysis

(expand state main findings)

0.25/0.5  
0.5/1

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

poor

satisfactory

good

excellent

10. Other comments:

- keep Table 1 Together

- refs for methods

- model fitting section incomplete

3.25

Don't need to re-do

logistic: Name

Medya Tekes Mizrakli

10.25/12 → 5.125/6

1. Formatting:

all margins 2.5cm

informative title

12 pt size

name on all pages

0.5/0.5

no raw R code or output

all pages numbered

max 7 pages

no blurry plots (NOT png)

6/6  
good job!

2. Introduction/Background:

brief statement of scientific question

all variables defined in EDA

0.5/0.5

3. EDA:

univariate numerical

bivariate numerical (cor)

univariate graphical

bivariate graphical

2/2

4. Model fitting:

give mathematical definition of model

state how model fitted (ie, maximum likelihood)

CLEARLY describe how model selected

Explain the steps

define all terms

1.75/2

5. Model assessment:

CLEARLY state model assumptions: + give PRIMARY reference

1. binary outcome

2. independent obs

3. linear relation between logit and linear predictor

4. no multicollinearity 5. no outliers (6. large sample size)

carry out assessment (numerical / graphics):

is this the right plot? scatterplots of logit vs. predictors (linearity assumption)

DEFINE -> Cook's distance / standardized residuals (outliers)

vif (to identify multicollinearity)

not clear

logit(p) (not logit( $\hat{p}$ ))

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

0.75

**hat** on response variable  
(ok if coefs in table)

max **2 sig digits** on coefs

7. Plots:

label size (not too small)

captions

(placement)

**NOT BLURRY**

8. Conclusions

(expand)  
recap analysis

(start new paragraph  
state main findings)

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

0.5/0.5

poor

satisfactory

good

excellent

0.75

10. Other comments:

- use single column

- primary refs methods

- good job!

4.25

Don't need to re-do

logistic: Name \_\_\_\_\_

Bettina Weber

10.5/12 → 5.25/6

6/6  
good job!

1. Formatting:

all margins 2.5cm

informative title

0.5/0.5

12 pt size

name on all pages

**no raw R code or output**

all pages numbered

~~max 7 pages~~

no blurry plots (**NOT** png)

0.5/0.5

2. Introduction/Background:

(OK)

brief statement of scientific question

A bit short...

all variables defined

2/2

3. EDA:

Don't need QQ

univariate numerical

bivariate numerical (cor)

univariate graphical

bivariate graphical

square

4. Model fitting:

1.5/2

give mathematical definition of model

state how model fitted (ie, maximum likelihood)

**CLEARLY** describe how model selected

Null, all hyps mathematically

define all terms

5. Model assessment:

1.5/2

**CLEARLY** state model assumptions: + give PRIMARY reference

1. binary outcome

2. independent obs

3. linear relation between logit and linear predictor

4. no multicollinearity 5. no outliers (6. large sample size)

carry out assessment (numerical / graphics):

scatterplots of logit vs. predictors (linearity assumption)

**DEFINE** -> Cook's distance / standardized residuals (outliers)

vif (to identify multicollinearity)

Define

$\hat{y} = P(\text{Surv} | \text{age, nodes})$  ← hat

1/1

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

hat on response variable  
(ok if coefs in table)

max **2 sig digits** on coefs

1/1

7. Plots:

(label size (not too small)

captions

placement

**NOT BLURRY**

1.25/1.5

8. Conclusions

recap analysis

(start new paragraph  
state main findings and expand)

0.5/0.5

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

poor      satisfactory      good      excellent

0.75/1

10. Other comments:

(give primary refs)

- good job! (just a few small things)

4.5

Don't need to redo

Maximilian Wettstein

10/12 → 8/6

logistic: Name \_\_\_\_\_

1. Formatting:

all margins 2.5cm

12 pt size

**no raw R code or output**

~~max 7 pages~~

informative title

name on all pages

all pages numbered

no blurry plots (**NOT png**)

→ 8/6  
→ 6/6  
good job!

0.5/0.5

no R formulas

2. Introduction/Background:

brief statement of scientific question (ok)

all variables defined

background a bit short

0.5/0.5

3. EDA:

univariate numerical

univariate graphical

bivariate numerical (cor)

bivariate graphical

layout: fig 2 description before fig - too much blank space

2/2

4. Model fitting:

give mathematical definition of model

state how model fitted (ie, maximum likelihood)

CLEARLY describe how model selected

define all terms

AIC → LOWER = better

→ not the reason we can use ML

1.5/2

5. Model assessment:

CLEARLY state model assumptions: + give PRIMARY reference

1. binary outcome

2. independent obs

3. linear relation between logit and linear predictor

4. no multicollinearity 5. no outliers (6. large sample size)

carry out assessment (numerical / graphics):

scatterplots of logit vs. predictors (linearity assumption)

DEFINE -> Cook's distance / standardized residuals (outliers)

vif (to identify multicollinearity)

1.5/2

0.75/

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

**hat** on response variable  
(ok if coefs in table)

max **2 sig digits** on coefs

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:

- use primary methods refs where possible  
[3]

- good job!



Don't need to re-do

logistic: Name \_\_\_\_\_

Ruiqi Yu

10.25/12 → 5.125/6

6/6  
good job!

1. Formatting:

all margins 2.5cm

informative title

0.5/0.5  
12 pt size

name on all pages

**no raw R code or output**

all pages numbered

max 7 pages

no blurry plots (**NOT png**)

2. Introduction/Background:

(OK)  
not much mentioned on used methodology

0.5/0.5

brief statement of scientific question

all variables defined

3. EDA:

2/2

univariate numerical

bivariate numerical (cor)

univariate graphical

bivariate graphical

4. Model fitting:

1.5/2

give mathematical definition of model

state how model fitted (ie, maximum likelihood)

**CLEARLY** describe how model selected

define all terms

5. Model assessment:

AIC, BIC, Pseudo R<sup>2</sup>

1.5/2

**CLEARLY** state model assumptions: + give **PRIMARY** reference

1. binary outcome

2. independent obs

3. linear relation between logit and linear predictor

4. no multicollinearity 5. no outliers (6. large sample size)

carry out assessment (numerical) / graphics):

scatterplots of logit vs. predictors (linearity assumption)

**DEFINE** -> Cook's distance / standardized residuals (outliers)

vif (to identify multicollinearity)

Define

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

0.5 / 1

**hat** on response variable  
(ok if coefs in table)

**max 2 sig digits** on coefs

7. Plots:

1 / 1

label size (not too small)

captions

placement

**NOT BLURRY**

8. Conclusions

1.5 / 1.5

recap analysis

state main findings

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

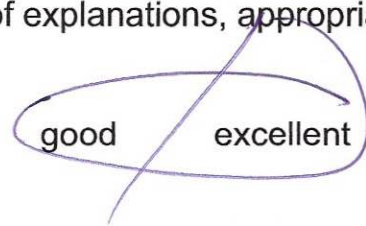
0.5 / 0.5

poor

satisfactory

good

excellent



10. Other comments:

0.75 / 1

- good job!

---

---

---

---

---

---

---

---

---

---

4.25

1. Formatting:

all margins 2.5cm

informative title

12 pt size

name on all pages

no raw R code or output

all pages numbered

max 7 pages

no blurry plots (NOT png)

2. Introduction/Background:

brief statement of scientific question

all variables defined

put in text, hard to read, no mention the types of vars

3. EDA:

not univariate 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

5. Model assessment:

not yet written

CLEARLY state model assumptions: + give PRIMARY reference

1. binary outcome

2. independent obs

3. linear relation between logit and linear predictor

4. no multicollinearity 5. no outliers (6. large sample size)

carry out assessment (numerical / graphics):

scatterplots of logit vs. predictors (linearity assumption)

DEFINE -> Cook's distance / standardized residuals (outliers)

vif (to identify multicollinearity)

3.25

0.5

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

**hat** on response variable  
(ok if coefs in table)

max **2 sig digits** on coefs

0.5

7. Plots:

label size (not too small)

captions

placement

**NOT BLURRY**

0/1.5

8. Conclusions

*not done*

recap analysis

state main findings

0.25

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

0.25

poor

satisfactory

good

excellent

*(incomplete)*

10. Other comments:

*refs - data + methods*

*- incomplete*

1.5

logistic: Name \_\_\_\_\_

Xingyue Zhang

6.25/12 → 3.125/6

1. Formatting:

all margins 2.5cm

informative title

0.5/0.5 12 pt size

name on all pages

no raw R code or output

all pages numbered

max 7 pages

no R formulas

no blurry plots (NOT png)

2. Introduction/Background:

0.5/0.5

brief statement of scientific question

all variables defined

0.5/2

3. EDA:

- Don't do logistic in EDA

did analysis on response var

univariate numerical

bivariate numerical (cor)

missing

univariate graphical

bivariate graphical

4. Model fitting:

0.75/2

give mathematical definition of model

BEFORE fitting

state how model fitted (ie, maximum likelihood)

CLEARLY describe how model selected

define all terms

Odds ratio, Hosmer-Lemeshow, Deviance, unclear

5. Model assessment:

0.75/2

mentioned in pieces rather than listing, hard to read

CLEARLY state model assumptions: + give PRIMARY reference

1. binary outcome

2. independent obs

3. linear relation between logit and linear predictor

4. no multicollinearity 5. no outliers (6. large sample size)

carry out assessment (numerical / graphics):

too short

scatterplots of logit vs. predictors (linearity assumption)

DEFINE -> Cook's distance / standardized residuals (outliers)

vif (to identify multicollinearity) report too short

→ First, select model, THEN assess

0.75/1

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

<sup>P(1-0)</sup>  
in wrong place  
hat on response variable  
(ok if coefs in table)

max **2 sig digits** on coefs

1/1

7. Plots:

( )  
label size (not too small)  
placement

captions

**NOT BLURRY**

0.75/1.5

8. Conclusions

(expand)  
recap analysis

(expand starting in new paragraph)  
state main findings

0.25/0.5

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

poor satisfactory good excellent

0.5/1

10. Other comments:

- primary refs for methods
- not clear why you are using these models - how did you get them?
- need to structure report in logical order:  
EDA ; THEN define model mathematically;  
THEN fitting / selection according to clearly specified procedure (s) and criteria ; THEN assessment on the selected model

3.25