GWAS: Name

1. Formatting:

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

2. Introduction/Background:

brief background and statement of scientific question

all variables defined

3. EDA / PCA:

relevant histograms / exploratory plots (NO BOXPLOTS)

EXPLAIN RELATION between PCs and population stratification

plot pc2 (y-axis) vs pc1 (x-axis)

4. Pre-processing / QC steps: CLEARLY EXPLAIN

SNP QC: criteria and reasons

sample QC: criteria and reasons

Hardy-Weinberg equilibrium: DEFINE and say how it relates to quality

measure of LD and how it is used

5. Association / post-association analysis: + give PRIMARY references

write out model *mathematically* (for a given SNP): CLEARLY define all variables and domains - Model MUST RELATE TO SNP

describe association analysis in words and mathematically

Manhattan plot and explanation

lambda analysis (*including* mathematical definition of lambda and <u>SQUARE</u> QQ plots)

6. Identify significant SNPs (include all relevant information)

7. Plots:

label size (not too small)	informative captions
placement	explanations
8. Conclusions	
1. recap analysis	2. state and interpret main findings

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

poor	satisfactory	good	excellent
		0	

10. Other comments:

A – no / incomplete / insufficient references

B - cite PRIMARY refs (not course notes, not wikipedia, etc.)

C - interpretation (cannot conclude causation, only association)

D - use your OWN WORDS / no apparently unattributed quotations

E – Intro: **1.** Give context; **2.** Clearly state scientific question; **3.** Describe data

- F univariate graphical: histograms not boxplots
- G (mathematical) model misspecified / unclear
- H clearly EXPLAIN / INTERPRET PLOTS (don't just state conclusions)
- I plot size / aspect ratio (make 'pretty')

Other: