

Lab 2 GSI grading

This form contains the final scores and extra comments from the GSI.

The respondent's email address (**rebeccabarter@berkeley.edu**) was recorded on submission of this form.

Name of student

Amy Ko

The student ID of the student whose paper you are grading *

24978168

Readability of report (5 points) *

	1	2	3	4	5	
Narrative unclear and/or difficult to read	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	Narrative very clear and/or easy to read

Grammar of report (5 points) *

	1	2	3	4	5	
Incorrect written grammar pervasive	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	Excellent written grammar

Analysis: redwood trees

In this section you will assess the actual analysis using kernel density estimation and loess on the redwood trees data.

Detail of kernel density estimation analysis (3 points) *

	0	1	2	3	
Did not explore different bandwidths or kernels	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	Explored a variety of bandwidths and kernels and clearly related these to the bias-variance-tradeoff

Relevance and quality of figures related to kernel density estimation (3 points) *

	0	1	2	3	
Did not provide any figures	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	Provided clear, relevant and visually appealing figures

Discuss one (or more) things that you liked about the author's kernel density estimation figures

Discuss one (or more) things that could be improved for the author's kernel density estimation figures

Detail of loess smoothing analysis (3 points) *

	0	1	2	3	
Did not conduct an analysis using a loess smoother	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	Explored a variety of bandwidths and polynomials and clearly related these to the bias-variance-tradeoff

Relevance and quality of figures related to loess smoothing (3 points) *

	0	1	2	3	
Did not provide any figures	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	Provided clear, relevant and visually appealing figures

Discuss one (or more) things that you liked about the author's loess figures

.....

Discuss one (or more) things that could be improved for the author's loess figures

.....

Analysis: linguistic survey

Level of detail in the written comparison between two questions (3 points) *

	1	2	3	
Little detail (barely described the relationships between the two questions)	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	Very detailed (described clearly the geographical groups formed by each question and discussed how the questions were related to one another)

Optional comments about author's analysis of the two questions

Quality and relevance of figures (e.g. maps) for the two questions (3 points)

*

	0	1	2	3	
Did not provide figures	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	Provided clear, informative, and visually appealing figures

Discuss one (or more) things that you liked about the author's figure(s)

Nice simple linked brushing app! Usually this kind of interactivity is more helpful when the axis on each plot are different, but this works well since you can isolate individual answers.

Discuss one (or more) things that could have been improved for the author's figure(s)

Discovered that the binary encoding should be aggregated (e.g. in lat-long bins) in order to perform meaningful PCA (or other dimensionality reduction technique) (2 points) *

	0	1	2	
Did not mention that dimensionality reduction did not work well on the binary encoded data	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	Found that PCA was ineffective for binary encoding and used aggregated data instead (e.g. grouped by ZIP or lat/long bins)

Discussed clustering and related these clustering results to geography (3 points) (note: deduct a point if the author used lat/long as a variable in their cluster algorithm) *

	0	1	2	3	
Did not discuss clustering	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	Discussed in detail the clusters found in the data and how they related to geography

Optional comments on cluster analysis

Quality and relevance of figures related to clustering and geography (3 points) *

	0	1	2	3	
No figures provided	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	Provided clear, informative, and visually appealing figures

Discuss one (or more) things you liked about the author's clustering figures

Discuss one (or more) things that could be improved for the author's clustering figures

Analyzed the robustness/stability of a finding (3 points) (give partial points if the author showed stability only by re-running K-means without perturbing the data) *

	0	1	2	3	
Did not study robustness	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	Tested in detail the robustness of their finding (e.g. using repeated data perturbations, subsamples, or bootstrapped samples)

Bonus point for a particularly cool visualization (i.e. not just scatter points on a map) (1 bonus point)

☐ The author made a really creative map!

Bonus point for a particularly cool analysis (i.e. answering a question of the data not required by the lab) (1 bonus point)

☒ The author performed a really creative analysis!

Reproducibility

In this section you will assess the reproducibility of the your peer's report. Be sure to take note of any extra README files that explain any extra steps you might need to take to recompile the report. If they have saved their figures in a separate folder, check to see whether there is a script that will automatically produce AND SAVE their figures. If not, take a point off for reproducibility.

Several people will have saved a large file (probably geocoded locations) and used this file in analysis. This is fine if they also provided clear instructions concerning how the reviewer could reproduce this file in an automated way (e.g. by running an R script or calling a function). If they rely on such a file but do not provide instructions about how one could reproduce this file, then take a point off for reproducibility. You do not need to actually regenerate this file.

Reproducibility of report (4 points) *

	1	2	3	4	
Could not recompile the report	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	Could recompile the report and figures without manual effort and got the same output as provided in the original pdf

If you could not recompile the report, or got different output, explain what went wrong

Readability of code (4 points) - be sure to look at any files in the R/ folder *

	1	2	3	4	
Code very difficult to read with little documentation	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	Code easy to read with clear documentation

Suggestions to improve code (either provide specific examples or general comments)

Clarity of folder structure (2 points) *

0

1

2

Many excess
files not relevant
to the report



The purpose of
each file is clear
and there are no
excess files in
the lab2 folder

Optional suggestions for improving folder structure

Concluding remarks

In this section you will provide some general feedback to the author.

One or more things that you liked about the report overall

One or more things that could be improved upon

Any other comments that you would like to add?

See your peers reviews for more detailed feedback

This form was created inside of UC Berkeley.

Google Forms