Lab 2 GSI

Name of student *

Mengling Liu

Student ID *

3033122406

Readability of report (5 points) *

	1	2	3	4	5	
Narrative unclear and/or difficult to read and/or there is not much detail in the explanations	◯	◯	⬤	◯	◯	Narrative very clear and/or easy to read and lots of detail is given

Grammar of report (5 points) *

	1	2	3	4	5	
Incorrect of written grammar pervasive	◯	◯	◯	⬤	◯	Excellent written grammar

Detail of kernel density estimation analysis (3 points) *

	0	1	2	3	
Did not explore different bandwidths or kernels	○	○	◉	○	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 *

	0	1	2	3	
Did not provide any figures	○	○	◉	○	Provided clear, relevant, and visually appealing figures

Comments on what you liked about the author's kernel density estimation figures *

I like how you labeled each plot

Comments on what could be improved in the author's kernel density estimation figures *

If you are combining so many plots into a single figure to avoid clutter include less tick marks on your y-axis

Detail of loess smoothing analysis (3 points) *

	0	1	2	3	
Did not conduct an analysis using a loess smoother	○	○	◉	○	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	○	○	◉	○	Provided clear, relevant, and visually appealing figures

Comments on what you liked about the author's loess figures *

Good job not just using the default theme!

Comments on what could be improved in the author's loess figures *

Use transparency, and maybe a little color point as the emphasis is on the trend line.

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)	●	○	○	Very detailed (described clearly the geographical groups formed by each question and discussed how the questions were related to one another)

Comments on authors 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	○	○	●	○	Provided clear, informative, and visually appealing figures

Comments on what you liked about the figure(s) *

I like that your legend has the actually answers to the questions

Comments on what could be improved in the author's figure(s) *

Because there are so many points that overlap transparency is necessary here. Additionally, a different color scheme would have been good because some of these color stick out more distinctly than others making it see like that answer is more important (i.e. the bright green)

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

	0	1	2	
Did not mention that dimensionality reduction did not work well on the binary encoded data	●	○	○	Found that PCA was ineffective for binary encoding and used aggregated data instead

Discussed clustering and related these clustering result to geography (3 points) NOTE: point subtracted if lat-long values included in cluster algorithm *

	0	1	2	3	
Did not discuss clustering	○	○	●	○	Discussed in the detail the clusters found in the data and how they relate to geography and thought carefully about number of clusters

Comments on clustering analysis

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

	0	1	2	3	
Did not provide figures	○	○	●	○	Provided clear, informative, and visually appealing figures

Comments on what you liked about the figure(s) related to clustering and geography *

Good choice of point size

Comments on what could be improved in the author's figure(s) related to clustering and geography *

Again, use transparency.

Also, make sure to have captions on all your figures

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

	0	1	2	3	
Did not study robustness	○	○	●	○	Tested in detail the robustness of their findings (e.e. 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 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 point)

☐ The author performed a creative analysis

Reproducibility of report (4 points) *

	1	2	3	4	
Could not recompile	◯	◉	◯	◯	Could recompile the report and figures without manual effort and got the same output

If you could not compile, or got different output, explain what went wrong

The data would not load

Readability of code (4 points) *

	1	2	3	4	
Code very difficult to read with little documentation	◯	◯	◉	◯	Code easy to read with clear documentation

Comments on code

A lot your lines go over 80 characters

Clarity of folder structure (2 points) *

	0	1	2	
Many excess files not relevant to report	○	○	⦿	The purpose of each file is clear and there are no excess files in the lab2 folder

Comments on what the author did well

Your analysis of PCA was very thorough!

Comments on what the author could improve

Make sure to give more detailed explanations of what you do. That will help me give you more points, because if you do not explain it I do not know why you made the decisions.

This form was created inside of UC Berkeley.

Google Forms