

STA 314H1S: Quiz 1

Time allowed: 5 minutes; Total points: 10

Student Name: _____ Student ID: _____

Please make sure your name and ID are also written clearly on your printout, and remember to hand in your printout together with the quiz. There will be 4 points for your printout. Answer the following question based on your code/output for the computation problem in the problem set 1.

1. (2 pts) For part(c), how will you modify your code to compute the training/testing **MAE**? **Circle** the line of source code that you will change in your printout, and **write the new code in the space provided below**. Will your conclusion change because of this modification? (*hint*: the function `abs` in R computes the absolute value of a vector)
2. (2 pts) For part(e), suppose now you want to modify the KNN regression algorithm so that the prediction is computed as a **median** instead of a mean of the neighbouring observations. How will you modify your code? **Circle** the line of source code that you will change in your printout, and **write the new code in the space provided below**. When some y_i 's are contaminated with measurement error, explain why the new algorithm could be better (*hint*: the function `median` in R computes the median of a vector).
3. (2 pts) For part(f), which K do you think has the best predictive performance? Which curve do you base your conclusion on? Explain your choice using no more than two sentences.