

STAT 500: Homework 1

Due Tuesday, Sep 19, 2017

Please **limit** your solution to at most 4 pages and **attach** your R-code.

1. Classify the following variables as binary, discrete, or continuous. Also classify them as qualitative or quantitative. Example: Age in years. Answer: Continuous and quantitative
 - (a) Time in terms of AM or PM.
 - (b) Angles as measured in degrees between 0 and 360.
 - (c) Bronze, Silver, and Gold medals as awarded at the Olympics.
 - (d) Height above sea level.
 - (e) Number of patients in a hospital.
2. Do exploratory data analysis for "Pima Indian" data.
 - (a) Find the dimension of the data (p and n).
 - (b) Find the numerical summaries of each variables (min, 1Q, median, mean, 3Q, max).
 - (c) Change some of the 0s to **NA** and Find the numerical summaries of each variables again.
 - (d) Which observation has the largest "diastolic blood pressure"? Give the row number.
 - (e) Draw the graphical summaries (histogram and boxplot) for a single variable "diastolic blood pressure". (**hint: you may want to use $par(mfrow = c(1,2))$**)
 - (f) Draw the graphical summaries (histogram and boxplot) for two variables; "test" and "bmi". You must present at least one graph showing the relationship between two variables. (**hint: you may want to use $par(mfrow = c(1,3))$ and $factor()$**)