

STAT-745

Homework 1:

Part I: Read Ch. 2 from ISLR and certify it in your submission. You should write:

"I certify that I indeed finished reading Ch. 2 from *An Introduction to Statistical Learning*, by James Gareth, Daniela Witten, Trevor Hastie, Robert Tibshirani"

assuming you indeed read that material.

Part II:

Perform the following tasks, which you should document in a brief report, including R code and results:

1. Read in the data from the object, which was saved with "save(galaxies, file="galaxies.RData")" R command.
2. Perform EDA of the data.
3. Fit a linear no-intercept model (called Hubble's Law):  $\text{velocity} = \beta_1 \cdot \text{distance} + \varepsilon$
4. Assess the quality of the model fit, but do not explore other models.
5. Estimate  $\beta_1$  (called Hubble's constant), including units. Hubble's constant is given in  $\text{km} \times \text{sec}^{-1} \times \text{Mpc}^{-1}$ . A mega-parsec (Mpc) is  $3.086 \times 10^{19}$  km. Velocity data is given in km per second, and distance in Mpc.
6. Find  $\beta_1^{-1}$  (which approximates the age of the universe) in seconds, and then transform it to years.