

## ST495/590 – Assignment 1 – Due 1/20

1. Read the paper “ESP paper rekindles discussion about statistics” and write a one-paragraph summary.
2. Load the ozone data from the course webpage. This dataset has daily ozone values for  $n = 1,106$  monitoring stations and  $m = 31$  days in July 2005.
  - (a) Create a table with the overall (across sites and days) mean, standard deviation, and percent missing.
  - (b) Write a loop to compute the mean, variance, and percent missing for each of the  $n$  sites; make a histogram of each variable (all three histograms should have  $n$  observations); create scatter plots of each pair of these variables (each of the three plots should have  $n$  points).
  - (c) Conduct a linear regression with response equal to the site’s mean and the site’s variance and percent missing as covariates

You should turn in your responses to these questions in 1-2 pages (i.e., one piece of paper with text on both sides). You should also turn in a separate file with carefully commented code. Only output in the 1-2 page document will be graded. Please staple both documents together!