IST772 Materials – Copyright 2018 by Jeffrey Stanton – Please do not post online.

## IST772 - Week 1 - Think Pair Share

Instructions: In your group or with your partner, discuss the question(s) below that were assigned to your group by your instructor. All students in the group should be prepared to share a response with the whole class.

- 1. What is meant by the term "statistical inference?" Describe the connection between the concept of "uncertainty" and statistical inference.
- 2. Describe the connections between the sum of squared deviations from the mean, the variance, and the standard deviation. Why do people like to discuss the standard deviation of a variable as opposed to the sum of squared deviations from the mean or the variance?
- 3. The mean of a very large data set of temperature readings is 72. The median is also 72. What, if anything, can you say about the shape of the distribution? What, if anything, can you say about the mode?
- 4. Annual temperature readings collected from Alpha City showed a mean of 72 (Fahrenheit) and a standard deviation of 6 degrees. Annual temperature readings collected from Omega City showed a mean of 72 (Fahrenheit) and a standard deviation of 18 degrees. In which city would you prefer to live and why?
- 5. A histogram shows a bell-shaped distribution containing exactly 1000 data points with a median of 120. How many data points have values higher than the median? If I draw a line at the 99th percentile, how many data points have values higher than that line? Assuming the distribution is symmetric, what is the mean?
- 6. In your new role as traffic analysts, you set up your measuring station alongside a small, but busy road where the speed limit is 30 MPH. Over the course of a day, you collect measurements of the speed of each vehicle that passes, the number of seconds between the arrival of each new vehicle at your location, and, through a cool new technology, the approximate weight of each vehicle. At the end of the day you look at histograms of all three variables. What do you see? Make sure to name the specific distribution that you believe might best model each variable.