1 Short Answer Questions

1. Briefly describe the two (acceptable) methods of data collection. Include in your description the advantages and disadvantages of each.

2. Briefly describe the difference between a matched pairs t-test and a two sample t-test. Include differences in the parameters being estimated in each case.

3. Indicate how you will be able to use techniques in this class in your job responsibilities.
2 Data Analysis

1. A common refrain in recent elections was turning communities into “green” communities. Fundamentally, communities are tasked with finding indicators for better air quality. Using the community indicators dataset, determine whether there is a significant linear relationship between air quality (max 90th percentile value) and Mean Travel Time to Work (Min.). Make sure to check whether the conditions for valid inference are met. If the conditions are not met, carry out the inference, but be sure to appropriately caveat your answers. Let $\alpha = 0.05$.

2. Using the productivity dataset, calculate a 95% confidence interval for the difference in mean productivity for employees with 10 years or less as an employees and employees with more than 10 years ($\mu_{0-10} - \mu_{11+}$). Note that you should check whether the conditions are met for the inference to be valid. State your conclusion in context of the problem. If the conditions are not met, carry out the inference, but be sure to appropriately caveat your answers. (HINT: This will require you to create a new variable.)

3. A common refrain in recent elections was turning communities into “green” communities. Fundamentally, communities are tasked with finding indicators for better air quality. Using the community indicators dataset, determine whether there is a significant linear relationship between air quality (max 90th percentile value) and Income. Make sure to check whether the conditions for valid inference are met. If the conditions are not met, carry out the inference, but be sure to appropriately caveat your answers. Let $\alpha = 0.05$. 
4. For the productivity dataset included on the Berman CD, compute a 90% confidence interval for the mean of Knowledge to perform job responsibilities. Make sure to check whether the conditions for valid inference are met. If the conditions are not met, carry out the inference, but be sure to appropriately caveat your answers. Include in your answer an appropriate statement of the confidence interval in context of the problem.

5. Using the employee attitudes dataset, is there a relationship between “Gender” and “The morale of Seminole County Government employees is high”? Why or why not? Include any graphs, numerical summaries to justify your answer (review steps discussed in class).

6. Using the employee attitudes dataset, compute a 95% confidence interval for the proportion of employees who Strongly Agree that “The work that I do is important”. Make sure to check whether the conditions for valid inference are met. If the conditions are not met, carry out the inference, but be sure to appropriately caveat your answers. Include in your answer an appropriate statement of the confidence interval in context of the problem.