BUS 735: Business Decision Making and Research
Instructor: Dr. James Murray
In-class Exercise: Bivariate Statistics

Learning Objectives:

- LO1: Construct and test hypotheses using a variety of bivariate statistical methods to compare characteristics between two populations.

- LO6: Be able to use standard computer packages such as R to conduct the quantitative analyses.

Directions: Work in groups of up to three people and answer the following questions. Type up your answers and submit a PDF document or Word Processing document that includes the relevant R output and upload your submission to the appropriate D2L Dropbox. For any questions that involve conducting a hypothesis test, be sure to do all of the following:

- Report what hypothesis test you are conducting

- State the null and alternative hypothesis

- State the p-value

- State the decision regarding rejection of the null hypothesis

- State the conclusion of the hypothesis test in plain English

Group assignment: Please include the names of all group members on your document. Please be sure that everyone is contributing and understanding everything.

Execute the following code to download and open the following data set concerning electricity prices:

```
load(url(''http://murraylax.org/datasets/electricity.RData''))
```

The data set includes average retail price of electricity (expressed in cents per kilowatthour) for residential customers, commercial customers, industrial customers, and overall for each of the lower 48 U.S. states in 2004 and 2005. The data set also includes the average temperature in each state, the disposable income per capita by state, and an indicator for the area of the country for each state (area = East, Midwest, South, and West).

1. Is there a difference in the average price for electricity between Eastern states and Midwestern states in 2005? If so, which area pays more? Report a 95% confidence interval for the average difference in price.

   R Hint: You can select a sub-sample with the `filter()` function as follows:

   `> filter(elec, Area=="Midwest" | Area=="East") %>% droplevels() -> elec.sub`

   The `droplevels()` function removes the other possibilities for levels of `Area` that are not in the sub-sample you create.

2. Is there a difference in the average price for electricity between residential and industrial customers in 2004? If so, who pays more? Report a 95% confidence interval for the average difference in price.

3. Is there a difference in the median average price for electricity between Eastern states and Midwestern states in 2005? Report the sample median for each and compute a 95% confidence interval for these estimates.

4. Is there a difference in the median average price for electricity between residential and industrial customers in 2004? Report the sample median for each and compute a 95% confidence interval for these estimates.