

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 in a 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

Everyone must type up and submit their own work. No copying and pasting from each other! Please put your name first, and all your group members names on your document. I will randomly select one submission from you group, evaluate that submission, and assign everyone the same grade. Please be sure that everyone is working together and understanding everything.

The dataset `electricity.csv` includes data on the average retail price of electricity (expressed in cents per kilowatt-hour) for residential customers, commercial customers, industrial customers, and overall for each of the lower 48 U.S. states in 2004 and 2005. The dataset also includes the average temperature in each state, the gross domestic product per capita (closely related to average income) 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 Help: To select the subset of the data that only includes Eastern and Midwestern states using the the following command:

```
elecsub <- elec[ elec$Area=='East' | elec$Area=='Midwest', ]
```

With the code above, `elecsub` will be a subset of the `elec` dataset only including *rows* where `Area` is equal to 'East' or 'Midwest'.

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.