BZAN 535 **Homework 5** **Due Sep. 28**

1. The carbo data lists two Kroger stores in zipcode 37075. This is Hendersonville, TN, just north of Nashville. The stores are #268 and #275.
    a) Write a query to capture the weekly carbo sales revenue for each of these stores. Show your query but not the resulting data. Export the data to JMP.
    b) Use Graph Builder to plot the weekly carbo sales revenue for the stores. The plot should show the data and a smoother for each store.
    c) What is the % increase in sales volume for Kroger (in Hendersonville) due to opening the second store? Give a single answer (a "point estimate") and explain how you computed this number. Also explain any assumptions upon which your answer depends.
    d) Did store #275 appear to suffer any loss in sales that coincided with the presence of this new store? (This might occur due to customers switching from #275 to #268.) Give a 95% confidence interval for the change in mean weekly sales. Explain in detail how you approached this problem. Finally, interpret the interval in the language of the problem.
    e) Was it necessary to adjust the standard error(s) in d) for autocorrelation? Justify your answer here.

2. Compare stores 229 and 321 in terms of carbo sales revenue. In particular, can we say from the available data that one store has higher sales on average than the other on an ongoing basis?
    a) Answer this question either with a 95% confidence interval for the difference in mean weekly sales or with a two-sided hypothesis test. (Was it necessary to adjust the standard error for autocorrelation? Justify your answer.)
    b) Also, comment about what assumptions you are making regarding this historical data when you make your assertion about any difference "on an ongoing basis."

3. Store #385 closed and a new store (#278) opened during the period of our data. Both stores are in Shelbyville, TN 37160.
    a) Plot weekly sales for these stores. Discuss what you see and how this makes answering 3b) a challenge. Also, what should we do with week 37 data? Why?
    b) Estimate the $ increase and the % increase in weekly sales based on this store closure/opening. Explain what data you used to answer this question and justify this choice.
    c) Construct a 95% confidence interval for the mean $ increase in weekly sales. Show how you calculated the interval. Be sure to take into account any autocorrelation if it is statistically significant.
    d) Comment about what might make the actual average weekly gain not coincide with our interval estimate.

4. This question involves computing an estimate for the expected weekly sales for Kroger in these four categories (in the available Kroger data).
    a) Show a query that computes the weekly total carbo sales and the number of Kroger stores with sales each week.
    b) Plot the data in an informative way and discuss what this reveals?
    c) JMP provides a 95% confidence interval of [$86,059, $89,520] for the expected Kroger weekly sales in these categories. Why is this interval not valid? That is, what assumption(s) is (are) violated?
    d) Provide an adjusted confidence interval. No need to interpret the interval, just show the computations.