

Problem Set #5: Statistical Analysis of Charlotte-Rock Hill Air Quality Data

Individual work required. Show all work to receive credit. You must provide both handwritten solved problems as well as a copy of your spread sheet with your submission

Download [dataset1](#) on the course syllabus page. The data set consists of ozone, NO, CO, and hydrocarbon ambient air measurements for several days during a recent ozone season in the Charlotte-Rock Hill Metropolitan Statistical Area.

Complete the following exercises:

1. For each of the variables (columns) listed in the data set, determine the mean, median, standard deviation, standard error, variance, CV, geometric mean, and range.
2. For CO readings, calculate the 90% confidence interval.
3. Use a normal statistical model from the data to predict the % of the time that the eight-hour ozone reading would be expected to be above 80 parts-per billion (ppb). Then compare this prediction with the actual % of readings that were above 80 ppb.
4. For the aromatics readings:
 - Assume that these data are normally distributed and then calculate the 95% confidence interval for the readings
 - From the two confidence intervals calculated above, which is a better representation of the data? Explain.