# ST433/533 – Mid-term Exam 1 – Due 9/23

The COVID-19 pandemic has affected most aspects of our lives, including our air quality. Your assignment is to perform a spatial analysis of the change in ambient PM2.5 concentration from 2019 to 2020 across the Southeastern US. To perform this analysis you must download data from

https://www.epa.gov/outdoor-air-quality-data/download-daily-data

For each station in VA, NC, SC, GA and FL, compute the difference between the average PM2.5 concentration for April-June 2020 and April-June 2019, i.e., $Y(\mathbf{s}) = \bar{Y}_{2020}(\mathbf{s}) - \bar{Y}_{2019}(\mathbf{s})$. Use only stations with at least 10 observations in both years. This is the response variable. You must also identity at least one covariate (other than simple functions of $\mathbf{s}$) you suspect may be predictive of this change (it will be helpful if this covariate is available at the monitors and the prediction locations, but other than this constraint, be creative!). Your objectives are to (i) map this difference across the southeast US, (ii) test for the significance of the covariate(s) you select and (iii) test for a statistically significant change in average PM2.5 at each location.

Summarize your analysis in a 3-6 page report (double spaced, 11pt, one-inch margins). Papers longer than six pages will be penalized. To avoid penalty, your report MUST have the follow sections and contents:

1. Introduction: Briefly describe the problem and your objectives

2. Data description: Describe and plot the data used in your analysis

3. Methods: Describe the model(s) you propose, with at least a couple of equations, and other relevant details

4. Model comparisons: Fit 3-5 models to the data, select a best fitting model and summarize how well we can predict using the final model

5. Model checking: List the main assumptions of your final model and verify that these assumptions are not blatantly violated

6. Summarize the final model: Give the estimates of all model parameters and interpret them in the context of the problem

7. Spatial prediction: Produce maps of the estimated difference between 2020 and 2019, the standard error of your estimates, and locations where the difference is statistically significant

8. Conclusions: Briefly summarize the main results of your paper

Your paper should be written as a professional document will full paragraphs, clearly labeled and numbered figures and/or tables, and few spelling/grammar errors. You should include sufficient detail that another student in class could reproduce your results. You do not need to turn in code. **This is an individual exam and you are not allowed to discuss the exam with other students or anyone other than the instructor. Doing so will have severe consequences.**

Email your completed exam to the instructor (bjreich@ncsu.edu) by midnight on the 23rd.

Have fun!