Stats 506 Group 16 Project Proposal

Team Members: Jaehoon Lee, Jiali Luan, Guanghui Min

Data Analysis Method:
We are going to use the Mixed Effect Model to analyze our dataset. The advantage of this model is that it can be used on data that are collected and summarized in groups.

Dataset:
We will be using the Beijing PM2.5 Data Set from the UCI Machine Learning Repository. Our response variable is going to be the PM2.5 index. We are still deciding on our fixed predictors and random predictors.

Languages:
We are planning to use Python and R to finish this data analysis project. For Python, we are going to adopt the package "Statsmodels", and for R we will be using "lme4" and "glmmADMB" packages.