

Capstone Proposal

Air quality is an issue that affects the livelihood of millions of people, especially those with respiratory issues, and at times it can affect the ability for people to spend time outdoors. In this report I will attempt to build a model to predict ozone concentrations in the air, based on hourly weather observations. Ozone is one of six pollutants whose atmospheric concentration is regularly tracked for determining air quality. Although ozone in the stratosphere plays a crucial role in protecting humans from the harmful effects of ultraviolet radiation from the sun, ozone near the ground is harmful for public health, especially among individuals with asthma or other respiratory issues. For this report, my clients would be health and wellness firms, companies that rely on summer tourism, and others who have a vested interest in maintaining a healthy air quality. If an effective model from this analysis can be used to reliably predict the ozone concentration, then my clients will benefit financially and potentially save people from the health problems associated with poor air quality.

This report will be using hourly data of ozone concentration from Essex, MD (a nearby eastern suburb of Baltimore), as well as weather data from Martin State Airport, which is located close to Essex. The air quality data was acquired from the Environmental Protection Agency (EPA) website, while the weather data was acquired from the National Oceanic and Atmospheric Administration (NOAA).

The approach to solving this problem will be to merge the datasets into one dataframe with ozone and weather data from 2006 to 2015, and then do an analysis to find a correlation and eventually attempt to build a model that can predict the ozone levels using weather data. The deliverables for this assignment will consist of Python code from the Jupyter notebook data containing data wrangling, data storytelling, inferential statistics, machine learning (which would include a predictive model), and the conclusions. This data will be incorporated into a PDF document, and a slide deck will be created for a presentation.