Womb for Concern: Air Quality Impact on Maternal Health Outcomes

Jade Glaister and Anja Vogt

Problem Statement

Project Description: We investigate the relationship between air quality index and maternal health outcome markers including infant birthweight and maternal eclampsia. The air quality data are from the EPA Outdoor Air Quality monitors, and maternal health data are Vital Statistics Natality Birth Data from the National Center for Health Statistics. We train classification models (Ridge and SVM) on natality data both with and without air quality index data to predict instances of maternal eclampsia and low infant birth weight. The models are evaluated using training and testing accuracies and rates of false negatives, which we aim to minimize. A highly accurate prediction model of this type can provide useful information to pregnant individuals about their risk level for eclampsia or infant low birth weight based on geographic, demographic, and health risk factors.

Data Gathering

Air Quality Data - [Air Quality Data Collected at Outdoor Monitors Across the US | US EPA](#)

- United States Environmental Protection Agency, *Outdoor Air Quality Data*
- Yearly summary of AQI values in a county: Includes maximum, 90th percentile, and median AQI, the count of days in each AQI category, and the count of days when the AQI could be attributed to each criteria pollutant
- Downloadable CSV file from website

Maternal and Infant Health Data - [Vital Statistics Natality Birth Data | NBER](#)

- National Center for Health Statistics, *Vital Statistics Natality Birth Data*.
- From the National Bureau of Economic Research website
- Includes demographic and health microdata for births occurring during the calendar year. Based on information abstracted from birth certificates filed in vital statistics offices of each US state and Washington, DC.
- Births from 2004, the most recent year that included county of birth occurrence

Defining Stakeholders

- Women and pregnant individuals
- Healthcare systems and professionals working in reproductive health, child and infant health

- World aid organizations, international climate aid organizations
- Government, policymakers, and regulatory bodies interested in the social, political, and economic implications of air quality and climate
- Health technology, biotech, pharmaceutical companies

Key Performance Indicators (KPIs)

Our main KPI is to minimize false negatives, such that the model minimizes the number of predicted normal birth weights and no eclampsia that are in reality low birth weights or instances of eclampsia. We will also investigate training accuracies, testing accuracies, and recall for more information about the model's performance.

References

Cleveland Clinic. Eclampsia. Last modified March 28, 2023. Accessed June 27, 2025. https://my.clevelandclinic.org/health/diseases/24333-eclampsia.

Liu, Anqi, Hui Li, Qian Zhang, et al. "Machine Learning Models for Predicting Low Birth Weight: A Population-Based Retrospective Cohort Study." BMC Pregnancy and Childbirth 24 (2024): Article 20. https://pubmed.ncbi.nlm.nih.gov/40566361/.

National Center for Health Statistics (*2025*). Data File Documentations, *Natality*, *2004.* National Center for Health Statistics, Hyattsville, Maryland. https://www.nber.org/research/data/vital-statistics-natality-birth-data.

United States Environmental Protection Agency. *Outdoor Air Quality Data*. Accessed June 27, 2025. https://www.epa.gov/outdoor-air-quality-data

World Health Organization. Low Birth Weight. Nutrition Landscape Information System. Accessed June 27, 2025. https://www.who.int/data/nutrition/nlis/info/low-birth-weight.

Wu, Jun, et al. "Residential Exposure to Traffic and Spontaneous Preterm Birth." Environmental Health Perspectives 117, no. 9 (2009): 1409–1413. https://doi.org/10.1289/ehp.0800334.

Disclaimer: Any analyses, interpretations, or conclusions reached are credited to Jade Glaister and Anja Vogt and not to NCHS, which is responsible only for the initial data.