Data used in this paper is all publicly available.

* 15-minute averaged PurpleAir data from all indoor and outdoor sensors was downloaded from PurpleAir’s ThingSpeak API.
  + Documentation: <https://api.purpleair.com/>
    - Contains data dictionary descriptors of each data field (created\_at, PM2.5\_CF\_1\_ug/m3\_A, PM2.5\_CF\_1\_ug/m3\_B, PM2.5\_CF\_ATM\_ug.m3\_A, PM2.5\_CF\_ATM\_ug.m3\_B, Temperature\_F\_A, Humidity\_%\_A, ID, Lat, Lon, Location, Type were the fields used in this analysis)
  + Time period: 1 January, 2017 to 20 July 2021
* The hourly averaged AQS data was downloaded from the AQS API
  + Documentation: <https://aqs.epa.gov/aqsweb/documents/data_api.html>
    - Contains data dictionary descriptors of each data field (Date GMT, Time GMT, State Name, County Name, Site Number, Parameter Name, Latitude, Longitude, Datum, Method Name, Sample Measurement were the fields used in this analysis)
  + Time period: 1 January, 2017 to 20 July 2021
  + Sensor monitor pairs were determined as those within 50 meters from a reference monitor
* For the code used to generate these datasets please contact the corresponding author
  + Priyanka DeSouza, Department of Urban and Regional Planning, University of Colorado Denver, Denver CO, USA, CU Population Center, University of Colorado Boulder, Boulder CO, USA
  + priyanka.desouza@ucdenver.edu