Data Dictionary - Dataset Information	
Dataset Name	Air Quality and Health Impacts
Dataset URL	https://data.cityofnewyork.us/Environment/Air-Quality/c3uy-2p5r
Data Provided by *The name of the NYC agency providing this data to the public.*	NYC Department of Health and Mental Hygiene
Each row is a... *The unit of analysis/level of aggregation of the dataset*	neighborhood average
Publishing Frequency *How often changed data is published to this dataset. For an automatically updated dataset, this is the frequency of that automation*	annually
Data Change Frequency *How often the data underlying this dataset is changed*	annually
Frequency Details *Additional details about the publishing or data change frequency, if needed*	
Dataset Description *Overview of the information this dataset contains, including overall context and definitions of key terms. This field may include links to supporting datasets, agency websites, or external resources for additional context.*	The NYC Environment & Health Data Portal shows how environments affect health, by publishing data and information that explain the connection. This dataset includes neighborhood averages of different air pollutants, estimates of emissions from buildings, traffic volumes and health impacts of air pollution. To explore this data alongside other NYC Environment and Health indicators, go to on.nyc.gov/dataportal
Why is this data collected? *Purpose behind the collection of this data, including any legal or policy requirements for this data by NYC Executive Order, Local Law, or other policy directive.*	The NYC Environment & Health Data Portal shows how environments affect health, by publishing data and information that explain the connection.

How is this data collected?	Each indicator has a separate data collection and analysis methodology.
The methods used to create and update this dataset, including what cleaning or processing was involved prior to dataset publication. *If data collection includes interpreting physical information this field includes technical details.* *If data collection includes fielding applications, requests, or complaints, this field includes details about the forms, applications, and processes used.*	Details can be found here: https://a816-dohbesp.nyc.gov/IndicatorPublic/data-explorer/
How can this data be used?	See on.nyc.gov/dataportal
Examples of and/or links to projects or agency operations that have used this dataset. *Where relevant, includes links to online projects, agency websites, visualizations, maps, or dashboards.* *What are some questions one might answer using this dataset?*	
What are the unique characteristics or limitations of this dataset?	
Unique characteristics of this dataset to be aware of, specifically, constraints or limitations to the use of the data.	
Additional geospatial information	Spatial files of boundaries used in this dataset can be downloaded here: https://github.com/nychealth/EHDP-data/tree/production/geography
For any datasets with geospatial data, specify the coordinate reference system or projection used and other relevant details.	

Data Dictionary - Column Information				
Column Name	Column Description	Expected/Allowed Values	Field Limitations	Additional Notes

Name of the column exactly as it appears in the dataset.	A brief, plain-language explanation of what the data in the column means.	Specifies if there is an expected range and/or format of possible values. For example, if the data type is Date & Time, this field will note whether the timestamp is MM/DD/YYYY or MM/YYYY. If the Column Name is ice cream, this field might note that values can be Chocolate, Vanilla or Strawberry. If relevant, this field specifies the unit of measurement of the data field, e.g. thousands, millions, $ value, miles, feet, year, etc.	Describes any unique characteristics or potential analytical limitations presented by this field, including: - the reasoning for any null, zero, or empty values in the data - if the data in the column was integrated from another dataset or organization - if the data covered includes a different time period - the source of the column and how the data in the column was generated. For example, information on how the data in this column was generated can include whether the data was self-reported directly by a person, system generated by	

			a database or agency system, derived through analytical manipulation of other fields or records; or obtained from a different agency.	
unique_id	Unique record identifier			
indicator_id	Identifier of the type of measured value across time and space		Indicators have the same data source, but may be broken into different measures	
name	Name of the indicator		Indicators have the same data source, but may be broken into different measures	
measure	How the indicator is measured		This could indicate that data values are rates or counts or percentages - as well as if the measure is a stratification of the indicator data.	

measure_info	Information (such as units) about the measure			
geo_type_name	Geography type	UHF' stands for United Hospital Fund neighborhoods	For instance, Citywide, Borough, and Community Districts are different geography types	
geo_join_id	Identifier of the neighborhood geographic area, used for joining to mapping geography files to make thematic maps		Reference the geo_type_name to know which geography files to join with	
geo_place_name	Neighborhood name		Could be 'Brownsville' or 'Brooklyn' or other designation depending on Geography type	
time_period	Description of the time that the data applies to		Could be a year, range of years, or season for example	
start_date	Date value for the start of the time_period		Always a date value; could be useful for plotting a time series	

data_value	The actual data value for this indicator, measure, place, and time			
message	notes that apply to the data value		For example, if an estimate is based on small numbers we will detail here	