

Data Dictionary:

Water Data:

Column Name	Description	Data Type	Example Values	Notes
BIN	Department of Buildings Building Identification Number. Unique ID for a building in NYC.	string	1234567	
BOROUGH	Borough where the water tank is located.	string	Manhattan	NYC has 5 boroughs: Manhattan, Brooklyn, Queens, Bronx, Staten Island.
ZIP	ZIP code of the water tank's location.	string	10001	Typically 5-digit ZIP codes.
HOUSE_NUM	House/street number of the water tank's location.	string	123	
STREET_NAME	Street name where the water tank is located.	string	Broadway	

BLOCK	Department of Buildings block number for the water tank's location.	string	123	Typically part of a unique identifier for a parcel.
LOT	Department of Buildings lot number for the water tank's location.	string	456	Typically part of a unique identifier for a parcel.
CONFIRMATION_NUM	Certification number provided by the Department of Health and Mental Hygiene (DOHMH).	string	C123456789	
REPORTING_YEAR	The year the inspection was reported.	integer	2023	Must be > 2000 and < current year.
TANK_NUM	Serial listing for multiple tanks at a given location.	integer	1	'First' tank is usually 1, increments for each additional tank.
INSPECTION_BY_FIRM	Name of the professional firm performing the inspection.	string	XYZ Engineering	
INSPECTION_PERFORMED	Whether the certifying inspection was performed.	string	Yes	Typically 'Yes' or 'No'.

INSPECTION_DATE	Date of inspection for the reported results.	date	01-05-2023	Format: YYYY-MM-DD.
GI_REQ_INTERNAL_STRUCTURE	Requirement for certifying inspection of the internal structure of the tank.	string	Yes	
GI_RESULT_INTERNAL_STRUCTURE	Result of the inspection of the internal structure of the tank.	string	N	N = No Unsanitary Conditions, A = Corrective Action Taken.
GI_REQ_EXTERNAL_STRUCTURE	Requirement for certifying inspection of the external structure of the tank.	string	Yes	
GI_RESULT_EXTERNAL_STRUCTURE	Result of the inspection of the external structure of the tank.	string	N	N = No Unsanitary Conditions, A = Corrective Action Taken.
GI_REQ_OVERFLOW_PIPES	Requirement for certifying inspection of overflow pipes.	string	Yes	
GI_RESULT_OVERFLOW_PIPES	Result of the inspection of overflow pipes.	string	N	N = No Unsanitary Conditions, A = Corrective Action Taken.

GI_REQ_ACCESS_LADDERS	Requirement for certifying inspection of access ladders.	string	Yes	
GI_RESULT_ACCESS_LADDERS	Result of the inspection of access ladders.	string	N	N = No Unsanitary Conditions, A = Corrective Action Taken.
GI_REQ_AIR_VENTS	Requirement for certifying inspection of air vents.	string	Yes	
GI_RESULT_AIR_VENTS	Result of the inspection of air vents.	string	N	N = No Unsanitary Conditions, A = Corrective Action Taken.
GI_REQ_ROOF_ACCESS	Requirement for certifying inspection of roof access.	string	Yes	
GI_RESULT_ROOF_ACCESS	Result of the inspection of roof access.	string	N	N = No Unsanitary Conditions, A = Corrective Action Taken.
SI_REQ_SEDIMENT	Requirement for certifying inspection for sediment in the tank.	string	Yes	

SI_RESULT_SEDIMENT	Result of the inspection for sediment in the tank.	string	N	N = No Unsanitary Conditions, A = Corrective Action Taken.
SI_REQ_BIOLOGICAL_GROWTH	Requirement for certifying inspection for biological growth in the tank.	string	Yes	
SI_RESULT_BIOLOGICAL_GROWTH	Result of the inspection for biological growth in the tank.	string	N	N = No Unsanitary Conditions, A = Corrective Action Taken.
SI_REQ_DEBRIS_INSECTS	Requirement for certifying inspection for debris or insects in the tank.	string	Yes	
SI_RESULT_DEBRIS_INSECTS	Result of the inspection for debris or insects in the tank.	string	N	N = No Unsanitary Conditions, A = Corrective Action Taken.
SI_REQ_RODENT_BIRD	Requirement for certifying inspection for evidence of rodents or birds.	string	Yes	

SI_RESULT_RODENT_BIRD	Result of the inspection for evidence of rodents or birds.	string	N	N = No Unsanitary Conditions, A = Corrective Action Taken.
SAMPLE_COLLECTED	Indicates if a water sample was collected for lab analysis.	string	Yes	
LAB_NAME	Name of the laboratory analyzing the water sample.	string	ABC Labs	
NYS_CERTIFIED	Confirms if the laboratory is certified by New York State.	string	Yes	Typically 'Yes' or 'No'.
ANALYTES	Describes the type of analysis performed on the sample.	string	B, O	B = Bacterial, O = Other.
COLIFORM	Certifying lab analysis results for total coliform.	string	A, P	A = Absent/None Detected, P = Present.
ECOLI	Certifying lab analysis results for E. coli.	string	A, P	A = Absent/None Detected, P = Present.
MEET_STANDARDS	Certifies if other analyses meet health standards.	string	Yes	

DELETED	Indicates if the record has been deleted or is inactive.	string	No	Generally 'Yes' or 'No'.
LATITUDE	Latitude of the water tank's location.	float	40.7128	
LONGITUDE	Longitude of the water tank's location.	float	-74.006	
COMMUNITY_BOARD	NYC Community Board District where the water tank is located.	string	12	NYC has 59 Community Boards.
COUNCIL_DISTRICT	NYC Council District where the water tank is located.	string	6	NYC has 51 Council Districts.
CENSUS_TRACT	U.S. Census Tract where the water tank is located.	string	123.45	Census Tracts are smaller geographic areas used for census data collection.
BBL	Borough, Block, and Lot (BB			

Air Data:

Column Name	Description	Data Type	Example Values	Notes
Date	Date when the data was recorded	date	09-05-2023	Format: YYYY-MM-DD

Source	Data source or organization that provided the data	string	EPA, AirQualityDept	Identifies the source of the data
Site ID	Unique identifier for the monitoring site	string	123456, 789012	May contain alphanumeric characters
POC	Point of Contact identifier for the monitoring site	integer	1, 2	Used to differentiate between multiple points at the same site
Daily Max 8-hour CO Concentration	Maximum 8-hour carbon monoxide concentration recorded at the site on the given day	float	4.5, 3.2	Measured in parts per million (ppm)
UNITS	Units used to measure the daily maximum 8-hour CO concentration	string	ppm, mg/m3	Typically 'ppm' for carbon monoxide
DAILY_AQI_VALUE	Daily Air Quality Index value calculated from the CO concentration	integer	75, 50	Represents the AQI value for the given day
Site Name	Name of the monitoring site	string	Central Park, Downtown Station	May vary based on naming conventions

DAILY_OBS_COUNT	Total number of observations taken during the day at the monitoring site	integer	24, 30	Represents the count of measurements
PERCENT_COMPLETE	Percentage of complete data recorded for the day at the monitoring site	float	100.0, 95.0	Indicates data completeness
AQS_PARAMETER_CODE	Parameter code used to identify the type of measurement	integer	42101	42101 typically refers to CO (Carbon Monoxide)
AQS_PARAMETER_DESC	Description of the AQS parameter	string	Carbon Monoxide	Describes the type of parameter
CBSA_CODE	Core-Based Statistical Area code for the region where the monitoring site is located	integer	35620	Identifies the CBSA region
CBSA_NAME	Name of the Core-Based Statistical Area	string	New York-Newark-Jersey City	Names may vary
STATE_CODE	Code for the state where the monitoring site is located	integer	36	Typically follows FIPS codes

STATE	Name of the state where the monitoring site is located	string	New York, California	
COUNTY_CODE	Code for the county where the monitoring site is located	integer	61	Typically follows FIPS codes
COUNTY	Name of the county where the monitoring site is located	string	New York County, Los Angeles County	
SITE_LATITUDE	Latitude of the monitoring site	float	40.7128, 34.0522	Coordinates for geographic location
SITE_LONGITUDE	Longitude of the monitoring site	float	-74.0060, -118.2437	Coordinates for geographic location