

Community Drinking Water

Environmental Public Health Tracking

AGGREGATE DATA SET SUMMARY

Drinking Water Quality

Characteristic	Description
Data Sources	Safe Drinking Water Act/State Drinking Water Information System (SDWA/SDWIS)
Purpose	<p>This data set contains the information needed to calculate Environmental Public Health Tracking (EPHT) measures of contaminants in public water supply for arsenic, disinfection byproducts, nitrates, atrazine, di(2-ethylhexyl) phthalate (DEHP), radium, tetrachloroethene (tetrachloroethylene) (PCE), trichloroethene (trichloroethylene (TCE), and uranium. Data are derived from state Safe Drinking Water Act databases. The data set consists of two tables:</p> <ol style="list-style-type: none"> 1. PWS Inventory. This file is required and contains descriptive and locational information about each public water system (PWS) with which water quality data is provided. This dataset should only include Community Water Systems (CWS) as defined and regulated by the Safe Drinking Water Act. It does not include Non-Transient Non-Community (NTNC) and Transient Non-Community water systems (TNC). There is one record for every year that a CWS was active, delivering drinking water to customers, and in which water quality data is complete. CWS that were once active and are currently inactive should be included if State's data support this scenario. 2. Drinking Water Quality Sampling Results. This file is required and contains one record for each community water system (CWS) for the mean and maximum concentrations per year of each of arsenic, disinfection byproducts, nitrates, atrazine, di(2-ethylhexyl) phthalate (DEHP), radium, tetrachloroethene (tetrachloroethylene) (PCE), trichloroethene (trichloroethylene (TCE), and uranium; and the mean concentrations per quarter of disinfection byproducts, nitrates and atrazine. This dataset also accommodates sample-level data and includes one record for each compliance sample for the same analytes used in calculating summary concentrations. Some fields (e.g. NumSamplingStations) are only included for summary-level data observations and belong to the schema group "SummaryLevelGroup". Other fields (e.g. DetectionLimit) are only included for sample-level data observations and belong to the schema group "SampleLevelGroup". At this time, sample-level observations are not required as part of the Community Drinking Water data submissions; associated optional fields for sample-level data are indicated as such in the Optionality column.
Restrictions	This is not a restricted access data set.

NATIONAL ENVIRONMENTAL PUBLIC HEALTH TRACKING
DATA DICTIONARY FOR AGGREGATE DATA
COMMUNITY DRINKING WATER
PUBLIC WATER SYSTEM INVENTORY

Field Name	Field Description	Data Type	Code Scheme and Legal values	Field Length	Optionality
StateFIPSCode	State FIPS code	String	FIPS State Code	2	Required
PWSIDNumber	PWS identifier	String	nine character value consisting of the 2 letter state abbreviation followed by 7 numbers	9	Required
YearAssociatedTo	Year that these data are associated with regards to sampling results	Text	YYYY. 1999 through latest complete year (e.g. 2011)	4	Required
YearPulled	Year that these data were pulled from state records	Text	YYYY. 1999 through latest year.	4	Required
PWSName	Name of PWS	Integer	Any; “U” = Unknown; “NS” = Not submitted	40	Required
PrincipalCountyServedFIPS	Principal county FIPS served by the CWS	String	FIPS County Code	5	Required
PrincipalCityFeatureID	Principal city, town or village Feature ID served by the CWS	Integer	9999999999;“-999” for Missing; “-888” for Not Submitted	10	Required

TotalConnections	Number of residential service connections	Integer	1-9999999”	7	Required
SystemPopulation	Permanent population uniquely served by the CWS	Integer	10-99999999 ”	8	Required
PrimarySourceCode	Type of source	String	GU = ground water under direct influence of surface water, GUP = purchased ground water under direct influence of surface water, GW = ground water, GWP = purchased ground water, SW = surface water, SWP = purchased surface water; “U” = Unknown; “NS” = Not submitted	3	Required
Latitude	Latitude in NAD83 decimal degrees describing approximate center of retail service area of water system	Integer	00.0000000 to 90.0000000;”-99.99” for Missing; “-88.88” for Not Submitted.	10	Required
Longitude	Longitude in NAD83 decimal degrees describing approximate center of retail service area of water system	Integer	-18.0000000 to 180.0000000;”-999” for Missing; “-888” for Not Submitted.	11	Required
LocationDerivationCode	Code describing how approximate latitude/longitude location was derived	String	SA = Service area polygon centroid; MFL = Mean of 1 or more facility locations that are expected to be proximate to service area extent; PCS = GNIS coordinates for Principal City Served; GSH = The geocoded address of water system headquarters; PNS - GNIS coordinates for Principal County Served; O= Other (e.g. zip code, etc.) “-999” = Missing; “-888” = Not Submitted;	3	Required

Drinking water quality sampling results

NATIONAL ENVIRONMENTAL PUBLIC HEALTH TRACKING
DATA DICTIONARY FOR AGGREGATE DATA
COMMUNITY DRINKING WATER
DRINKING WATER SAMPLING RESULTS

Field Name	Field Description	Data Type	Code Scheme and Legal values	Field Length	Optionality
StateFIPSCode	State FIPS code	String	FIPS State Code	2	Required
PWSIDNumber	PWS identifier	String	Nine character value consisting of the 2 letter state abbreviation followed by 7 numbers	9	Required
Year	Year	Text	YYYY; 1999 through latest complete year (e.g. 2011)	4	Required
AnalyteCode	USEPA Analyte code for required constituents (arsenic, nitrate, TTHM, HAA5, atrazine, PCE, TCE, DEHP, radium, and uranium)	Integer	1005=Arsenic; 2050=Atrazine; 2456=HAA5; 2950=TTHM; 2039=DEHP; 1040=Nitrate; 2987=PCE; 2984=TCE; 4010=Combined Radium 226 & 228; 4006=Uranium	4	Required
ConcentrationUnits	The analyte-specific units of summary-level measures and individual sample values as reported in the Concentration and DetectionLimit fields. Each analyte has a standard unit for this dataset.	String	“ug/L” allowed only for (Arsenic, TTHM, HAA5, Atrazine, DEHP, PCE,TCE, uranium); “mg/L” allowed only for (Nitrate as nitrogen); “pCi/L” allowed only for (Radium)	6	Required

NATIONAL ENVIRONMENTAL PUBLIC HEALTH TRACKING

DATA DICTIONARY FOR AGGREGATE DATA

COMMUNITY DRINKING WATER

DRINKING WATER SAMPLING RESULTS

Field Name	Field Description	Data Type	Code Scheme and Legal values	Field Length	Optionality
Concentration	Reported summary-level concentration or reported concentration of sample	Integer	>0 for summary-level measure or sample-level concentration; -888, if sample-level data and NonDetectFlag=1	6.4	Required
DateSampled	Date of sample (sample-level data) or Date last sampled (summary-level data)	Date	A valid date from 1/1/1999 through December 31 st of the latest complete year (e.g. 2011-12-31).	YYYY-MM-DD	Required
SamplePointID	Sampling station identifier for sample-level records only.	String	Character ID for sampling station;“-999” for missing; “-888” for Not Submitted	20	Optional
DetectionLimit	Sample detection limit	Integer	>0, if NonDetectFlag=1; -888, if NonDetectFlag=0.	3.6	Optional
NonDetectFlag	Flag to indicate whether sample resulted in a detection or not	Integer	1=Sample was a non-detect; 0=Sample was a detection	1	Optional

NATIONAL ENVIRONMENTAL PUBLIC HEALTH TRACKING

DATA DICTIONARY FOR AGGREGATE DATA

COMMUNITY DRINKING WATER

DRINKING WATER SAMPLING RESULTS

Field Name	Field Description	Data Type	Code Scheme and Legal values	Field Length	Optionality
AggregationType	The type of summary operation performed (i.e. mean or max) for summary-level data.	String	“X” = Mean (for Annual and Quarterly data); “MX” = Maximum (For Annual Data ONLY; DO NOT SUBMIT FOR QUARTERLY DATA)	3	Required
NumSamplingLocations	Number of compliance sampling locations available from which summary-level records were derived.	Required	1-9999; “-888” for Not Submitted	4	Required
SummaryTimePeriod	Year or Quarter for summary-level data	String	YYYY for annual summarized values; YYYY-Q for quarterly summarized values Allowed Only for AnalyteCodes 2050, 2456, 2950 and 1040	10	Required
NumSamples	The number of samples that were used in calculating the mean/max for a given analyte during a quarter or year.	Integer	1-XXXX	4	Required
NumNonDetects	The number of samples that were non-detections for summary-level data.	Integer	0-XXXX (XXXX must be no greater than NumSamples)	4	Required