



Data format for Project Tycho 2.0 dataset CSV files

2017-08-16

Project Tycho 2.0 datasets include counts of cases for conditions (i.e. disorders) in populations reported by health agencies. These counts can be specified for different geographical regions, for different time intervals, for various age ranges and/or subpopulations, and can represent fatalities (deaths) or not. Each datafile includes these, and other, attributes of counts, as indicated below.

Missing data: Counts are often not reported for all time intervals and locations. Project Tycho 2.0 datasets only include time intervals and location with counts available and users will have to insert time intervals for which no counts were reported to complete time series of counts. Project Tycho 2.0 uses the value “NA” for missing information about count attributes.

Project Tycho 2.0 dataset variables

Attribute (* required)	Description	Value allowed	URL of vocabulary for attribute values
ConditionName*	Name of reported condition as listed in SNOMED-CT	Any SNOMED-CT disorder name	https://biosharing.org/bsg-s000098
ConditionSNOMED*	SNOMED-CT code for reported condition (https://biosharing.org/bsg-s000098)	Any SNOMED-CT disorder code	https://biosharing.org/bsg-s000098
PathogenName	Organism name for pathogen causing reported condition	Any organism name in the NCBI Taxonomy, or NA	https://biosharing.org/bsg-s000154
PathogenTaxonID	NCBI Taxonomy identifier for pathogen causing reported condition	Any NCBI Taxonomy ID, or NA	https://biosharing.org/bsg-s000154
Fatalities*	Counts of reported condition represent fatalities	1, 0	No external standard
CountryName*	ISO 3166 English Short Name of country for which a count was reported	Any ISO 3166 country English Short Name (all capitals), as of 07-24-2017	https://www.iso.org/iso-3166-country-codes.html
CountryISO	ISO 3166 2-letter code for country	Any 2-letter ISO 3166 country code, as of 07-24-2017	https://www.iso.org/iso-3166-country-codes.html

Attribute (* required)	Description	Value allowed	URL of vocabulary for attribute values
Admin1Name	ISO 3166 Name of first administrative subdivision for which a count was reported	Any ISO 3166 first administrative subdivision name (all capitals), as of 07-24-2017, or NA	https://www.iso.org/iso-3166-country-codes.html
Admin1ISO	ISO 3166-2 code for first administrative subdivision	Any ISO 3166-2 code for first administrative subdivision, as of 07-24-2017, or NA	https://www.iso.org/iso-3166-country-codes.html
Admin2Name	Placename from Geonames of second-order administrative division for which a count was reported	Any placename from Geonames of second-order administrative division (all capitals), or NA	http://www.geonames.org/
CityName	Name of Populated place for which a count was reported, from Geonames	Any name from Geonames for (part of) a populated place, seat, capital, or similar feature class (all capitals), or NA	http://www.geonames.org
PeriodStartDate*	Start date of time interval for which a count was reported, in ISO 8601 format	Any date in ISO 8601 format	https://www.iso.org/iso-8601-date-and-time-format.html
PeriodEndDate*	End date of time interval for which a count was reported, in ISO 8601 format	Any date in ISO 8601 format	https://www.iso.org/iso-8601-date-and-time-format.html
PartOfCumulativeCountSeries*	Count is part of a series of cumulative counts (instead of being part of a series of fixed time interval counts)	0,1	No external standard
AgeRange*	Age range for which a count was reported (in years)	Any age range between 0 and 130 with age values separated by "-" as in: "0-130"	No external standard
Subpopulation*	A subgroup of the population for which a count was reported	"Civilian", "Military", "None specified"	No external standard
PlaceOfAcquisition	Place where reported condition was acquired: "Domestic" (i.e., in the country for which a case count was	"Domestic", "Abroad", or NA	No external standard

Attribute (* required)	Description	Value allowed	URL of vocabulary for attribute values
	reported) or “Abroad” (i.e., not in the country for which the case count was reported)		
DiagnosisCertainty	Qualifier for certainty of diagnosis for a count condition	“Definite”, “Equivocal”, “Possible diagnosis”, “Probable diagnosis”, or NA	http://purl.bioontology.org/ontology/SNOME/DCT/106230009
SourceName*	Name of the source (system, database, institution) from which counts were obtained by the Project Tycho team	E.g., “Nationally Notifiable Disease Surveillance System”, “World Health Organization DengueNet Database”	No external standard
CountValue*	The count value	Any whole number	No external standard

Example of part of a dataset

"ConditionName", "ConditionSNOMED", "PathogenName", "PathogenTaxonID", "Fatalities", "CountryName", "CountryISO", "Admin1Name", "Admin1ISO", "Admin2Name", "CityName", "PeriodStartDate", "PeriodEndDate", "PartOfCumulativeCountSeries", "AgeRange", "Subpopulation", "PlaceOfAcquisition", "DiagnosisCertainty", "SourceName", "CountValue" "Measles",14189004,"Measles virus",11234,0,"UNITED STATES OF AMERICA","US","WISCONSIN","US-WI",NA,NA,1951-12-30,1952-01-05,0,"0-130","None specified",NA,NA,"United States Morbidity and Mortality Weekly Report",191 "Measles",14189004,"Measles virus",11234,0,"UNITED STATES OF AMERICA","US","WISCONSIN","US-WI",NA,NA,1952-01-06,1952-01-12,0,"0-130","None specified",NA,NA,"United States Morbidity and Mortality Weekly Report",386 "Measles",14189004,"Measles virus",11234,0,"UNITED STATES OF AMERICA","US","WISCONSIN","US-WI",NA,NA,1952-01-13,1952-01-19,0,"0-130","None specified",NA,NA,"United States Morbidity and Mortality Weekly Report",460 "Measles",14189004,"Measles virus",11234,0,"UNITED STATES OF AMERICA","US","WISCONSIN","US-WI",NA,NA,1952-01-20,1952-01-26,0,"0-130","None specified",NA,NA,"United States Morbidity and Mortality Weekly Report",316 "Measles",14189004,"Measles virus",11234,0,"UNITED STATES OF AMERICA","US","WISCONSIN","US-WI",NA,NA,1952-01-27,1952-02-02,0,"0-130","None specified",NA,NA,"United States Morbidity and Mortality Weekly Report",506 "Measles",14189004,"Measles virus",11234,0,"UNITED STATES OF AMERICA","US","WISCONSIN","US-WI",NA,NA,1952-02-03,1952-02-09,0,"0-130","None specified",NA,NA,"United States Morbidity and Mortality Weekly Report",600 "Measles",14189004,"Measles virus",11234,0,"UNITED STATES OF AMERICA","US","WISCONSIN","US-WI",NA,NA,1952-02-10,1952-02-16,0,"0-130","None specified",NA,NA,"United States Morbidity and Mortality Weekly Report",479 "Measles",14189004,"Measles virus",11234,0,"UNITED STATES OF AMERICA","US","WISCONSIN","US-WI",NA,NA,1952-02-17,1952-02-23,0,"0-130","None specified",NA,NA,"United States Morbidity and Mortality Weekly Report",883 "Measles",14189004,"Measles virus",11234,0,"UNITED STATES OF AMERICA","US","WISCONSIN","US-WI",NA,NA,1952-03-02,1952-03-08,0,"0-130","None specified",NA,NA,"United States Morbidity and Mortality Weekly Report",887 "Measles",14189004,"Measles virus",11234,0,"UNITED STATES OF AMERICA","US","WISCONSIN","US-WI",NA,NA,1952-03-09,1952-03-15,0,"0-130","None specified",NA,NA,"United States Morbidity and Mortality Weekly Report",1121 "Measles",14189004,"Measles virus",11234,0,"UNITED STATES OF AMERICA","US","WISCONSIN","US-WI",NA,NA,1952-03-16,1952-03-22,0,"0-130","None specified",NA,NA,"United States Morbidity and Mortality Weekly Report",1211 "Measles",14189004,"Measles virus",11234,0,"UNITED STATES OF AMERICA","US","WISCONSIN","US-WI",NA,NA,1952-03-23,1952-03-29,0,"0-130","None specified",NA,NA,"United States Morbidity and Mortality Weekly Report",1456 "Measles",14189004,"Measles virus",11234,0,"UNITED STATES OF AMERICA","US","WISCONSIN","US-WI",NA,NA,1952-03-30,1952-04-05,0,"0-130","None specified",NA,NA,"United States Morbidity and Mortality Weekly Report",1463 "Measles",14189004,"Measles virus",11234,0,"UNITED STATES OF AMERICA","US","WISCONSIN","US-WI",NA,NA,1952-04-06,1952-04-12,0,"0-130","None specified",NA,NA,"United States Morbidity and Mortality Weekly Report",1467 "Measles",14189004,"Measles virus",11234,0,"UNITED STATES OF AMERICA","US","WISCONSIN","US-WI",NA,NA,1952-04-13,1952-04-19,0,"0-130","None specified",NA,NA,"United States Morbidity and Mortality Weekly Report",1844 "Measles",14189004,"Measles virus",11234,0,"UNITED STATES OF AMERICA","US","WISCONSIN","US-WI",NA,NA,1952-04-20,1952-04-26,0,"0-130","None specified",NA,NA,"United States Morbidity and Mortality Weekly Report",1958 "Measles",14189004,"Measles virus",11234,0,"UNITED STATES OF AMERICA","US","WISCONSIN","US-WI",NA,NA,1952-05-04,1952-05-10,0,"0-130","None specified",NA,NA,"United States Morbidity and Mortality Weekly Report",2464 "Measles",14189004,"Measles virus",11234,0,"UNITED STATES OF AMERICA","US","WISCONSIN","US-WI",NA,NA,1952-05-11,1952-05-17,0,"0-130","None specified",NA,NA,"United States Morbidity and Mortality Weekly Report",1482 "Measles",14189004,"Measles virus",11234,0,"UNITED STATES OF AMERICA","US","WISCONSIN","US-WI",NA,NA,1952-05-18,1952-05-24,0,"0-130","None specified",NA,NA,"United States Morbidity and Mortality Weekly Report",5225 "Measles",14189004,"Measles virus",11234,0,"UNITED STATES OF AMERICA","US","WISCONSIN","US-WI",NA,NA,1952-05-25,1952-05-31,0,"0-130","None specified",NA,NA,"United States Morbidity and Mortality Weekly Report",2475
--

In case of questions, please visit www.tycho.pitt.edu or email tycho@phdl.pitt.edu