



## Data format for custom-compiled Project Tycho 2.0 dataset CSV files

2018-03-28

Project Tycho 2.0 custom-compiled 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. Custom-compiled datasets also include for each count the DOI of the pre-compiled dataset that contains the count.

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 variables for custom-compiled datasets

Attribute (* required)	Description	Value allowed	Identifier of vocabulary for attribute values
ConditionName*	Name of reported condition as listed in SNOMED-CT	Any SNOMED-CT disorder name	doi:10.25504/fairsharing.d88s6e
ConditionSNOMED*	SNOMED-CT code for reported condition ( <a href="https://biosharing.org/bsg-s000098">https://biosharing.org/bsg-s000098</a> )	Any SNOMED-CT disorder code	doi:10.25504/fairsharing.d88s6e
PathogenName	Organism name for pathogen causing reported condition	Any organism name in the NCBI Taxonomy, or NA	doi:10.25504/fairsharing.fj07xj
PathogenTaxonID	NCBI Taxonomy identifier for pathogen causing reported condition	Any NCBI Taxonomy ID, or NA	doi:10.25504/fairsharing.fj07xj
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	<a href="https://www.iso.org/standard/63545.html">https://www.iso.org/standard/63545.html</a>

<b>Attribute (* required)</b>	<b>Description</b>	<b>Value allowed</b>	<b>Identifier of vocabulary for attribute values</b>
CountryISO	ISO 3166 2-letter code for country	Any 2-letter ISO 3166 country code, as of 07-24-2017	<a href="https://www.iso.org/standard/63545.html">https://www.iso.org/standard/63545.html</a>
Admin1Name	ISO 3166 Name of first administrative subdivision for which a count was reported	Any ISO 3166 fist administrative subdivision name (all capitals), as of 07-24-2017, or NA	<a href="https://www.iso.org/standard/63545.html">https://www.iso.org/standard/63545.html</a>
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	<a href="https://www.iso.org/standard/63545.html">https://www.iso.org/standard/63545.html</a>
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	<a href="http://www.geonames.org/">http://www.geonames.org/</a>
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	<a href="http://www.geonames.org/">http://www.geonames.org/</a>
PeriodStartDate*	Start date of time interval for which a count was reported, in ISO 8601 format	Any date in ISO 8601 format	<a href="https://www.iso.org/iso-8601-date-and-time-format.html">https://www.iso.org/iso-8601-date-and-time-format.html</a>
PeriodEndDate*	End date of time interval for which a count was reported, in ISO 8601 format	Any date in ISO 8601 format	<a href="https://www.iso.org/iso-8601-date-and-time-format.html">https://www.iso.org/iso-8601-date-and-time-format.html</a>
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

Attribute (* required)	Description	Value allowed	Identifier of vocabulary for attribute values
PlaceOfAquisition	Place where reported condition was acquired: "Domestic" (i.e., in the country for which a case count was reported) or "Abroad" (i.e., not in the country for which the case count was reported)	"Domestic", "Abroad", or NA	No external standard
DiagnosisCertainty	Qualifier for certainty of diagnosis for a count condition	"Definite", "Equivocal", "Possible diagnosis", "Probable diagnosis", or NA	doi:10.25504/fairsharing.d88s6e
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
DOI	Digital object identifier of the pre-compiled dataset that includes the count	Any Project Tycho DOI	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,PlaceOfAquisition,DiagnosisCertainty,SourceName,CountValue,DOIDengue hemorrhagic fever,20927009,Dengue virus,12637,0,LAO PEOPLE'S DEMOCRATIC REPUBLIC,LA,ATTAPU,LA-AT,NA,NA,2002-01-01,2002-01-31,0,0-130,None specified,NA,NA,World Health Organization DengueNet Database,0.0,10.25337/T7/ptycho.v2.0/LA.20927009Dengue hemorrhagic fever,20927009,Dengue virus,12637,0,LAO PEOPLE'S DEMOCRATIC REPUBLIC,LA,ATTAPU,LA-AT,NA,NA,2002-02-01,2002-02-28,0,0-130,None specified,NA,NA,World Health Organization DengueNet Database,0.0,10.25337/T7/ptycho.v2.0/LA.20927009Dengue hemorrhagic fever,20927009,Dengue virus,12637,0,LAO PEOPLE'S DEMOCRATIC REPUBLIC,LA,ATTAPU,LA-AT,NA,NA,2002-03-01,2002-03-31,0,0-130,None specified,NA,NA,World Health Organization DengueNet Database,0.0,10.25337/T7/ptycho.v2.0/LA.20927009Dengue hemorrhagic fever,20927009,Dengue virus,12637,0,LAO PEOPLE'S DEMOCRATIC REPUBLIC,LA,ATTAPU,LA-AT,NA,NA,2002-04-01,2002-04-30,0,0-130,None specified,NA,NA,World Health Organization DengueNet Database,0.0,10.25337/T7/ptycho.v2.0/LA.20927009Dengue hemorrhagic fever,20927009,Dengue virus,12637,0,LAO PEOPLE'S DEMOCRATIC REPUBLIC,LA,ATTAPU,LA-AT,NA,NA,2002-05-01,2002-05-31,0,0-130,None specified,NA,NA,World Health Organization DengueNet Database,0.0,10.25337/T7/ptycho.v2.0/LA.20927009Dengue hemorrhagic fever,20927009,Dengue virus,12637,0,LAO PEOPLE'S DEMOCRATIC REPUBLIC,LA,ATTAPU,LA-AT,NA,NA,2002-06-01,2002-06-30,0,0-130,None specified,NA,NA,World Health Organization DengueNet Database,0.0,10.25337/T7/ptycho.v2.0/LA.20927009Dengue hemorrhagic fever,20927009,Dengue virus,12637,0,LAO PEOPLE'S DEMOCRATIC REPUBLIC,LA,ATTAPU,LA-AT,NA,NA,2002-07-01,2002-07-31,0,0-130,None specified,NA,NA,World Health Organization DengueNet Database,0.0,10.25337/T7/ptycho.v2.0/LA.20927009Dengue hemorrhagic
```

In case of questions, please visit [www.tycho.pitt.edu](http://www.tycho.pitt.edu) or email [tycho@phdl.pitt.edu](mailto:tycho@phdl.pitt.edu)