# Standard Data Properties
## of
# Department Core Data Items


# Georgia Department of Public Health
# Office of Health Indicators for Planning
### First released July 2002

**Architect: Frank Millard, M.A**
**Developers:  Gordon R. Freymann, M.P.H., Michael A. Coletta, M.P.H., Rajni Samavedam, M.P.H., Cathie McCabe, M.B.A., Jacqueline R. Bennett-Burkholder, Ph.D., M.S., Robert Attaway**
**Project Lead:  Gordon R. Freymann, M.P.H.**

Definition of Terms

| Property | Definition |
| --- | --- |
| Variable Name (s) | Name of the data item used for storage and if applicable, presentation. Storage Name(s) for gold standard data items (variables) begin with GS. (Gold standard), followed by a domain identifier (e.g. .birth.), followed by an owner (e.g. .infant.), followed by the variable name (gs.birth.infant.sex).  Presentation Name(s) (or "label") is used when to present data, such as "mother's education level." |
| Definition and Variable Associated Standards | A statement containing the reason to collect or use the variable and external standards that applies to the variable. |
| Valid Values | Acceptable values for the variable being defined (e.g., mother's age range = 10-55 years inclusive.). |
| Categorical Attributes | Standard categories to be used for a variable when applicable (e.g. birthweight can be categorized into 'low birth weight' and 'very low birthweight').  Categories are groupings of valid values. |
| Data Type | The characteristic of a variable that determines what kind of data it can hold. For example, data types include Byte, Boolean, Integer, Long Integer, Currency, Decimal, String, Double, and Date. |
| Field Length | The number of numerical places or characters for a specific field. This is character field length, not data storage field length. |
| Unit of Measurement | Refers to the specific unit within a measurement system, at which measurements for a variable are made. Examples include birthweights measured in grams, or dates measured in days. "Unitless" refers to cases where a unit per se is not measured. |
| Level of Measurement | Nominal - categorical properties or labels; such as the variable race; Ordinal - objects are ordered by some nominal category irrespective of magnitude, and irrespective to the distance between ordered levels; such as a variable representing the level of agreement with a statement represented by disagree, somewhat agree, agree, disagree; Interval - ordering of objects is respective to a nominal category, the distance between objects respective to the nominal category, and without respect to the magnitude of the nominal category such as the number of hours a client waited for service measured in hours represented by 1, 2, 3, 4 etc.; and Ratio - objects are ordered respective to a nominal category, where the distance between objects is known, and each objects measurement is respective to a known zero value such as annual income measured in dollars represented by 20,051, 55,987, 42,042, etc. |
| Unit of Analysis. | The unit of measurement assigned to a variable for analysis. "Unitless" refers to cases where a unit per se is not analyzed. |
| Level of Analysis | The level of measurement assigned to a variable for analysis |

| Property | Definition |
| --- | --- |
| Derivation | For calculated fields, the data items used and method to derive the calculated variable. |
| Time Stamp of Standard | The date on which the variable definition was defined or revised. |
| Access | Some data items can by themselves or with other data items reveal the identity of a person and thus make record protected health information per HIPAA guidelines.  Therefore these data items are not public use.<br>Data items with demonstrated business value to the Department for health planning are considered datamart data items. |

PERSON (Age, Race, Ethnicity, Sex)

| Property | Value |
|---|---|
| Presentation Name(s) | AGE |
| Definition | Elapsed time since birth.  Age in years at the time of this event. Age is reported as age at last birthday - that is, age in completed years, and calculated by subtracting date of delivery from the reference date, with the reference date being the date of the examination, interview, or other contact with an individual (NCHS). |
| Valid Values | 0 - 44,194 days.  0 – 120 years.  (44,194 translates into 120 years); -1=unknown. Four (4) years = (365 x 4) + 1. |
| Data Type | Long Integer |
| Field Length | 5 |
| Unit of Measurement | Day |
| Level of Measurement | Interval |
| Unit of Analysis | Day |
| Level of Analysis | Interval |
| Derivation | EVENT_DATE –DOB = AGE IN DAYS.  CONVERT TO YEARS FOR PRESENTATION. |
| Time Stamp of Standard | 3/27/2002. |

| Property | Value |
|---|---|
| Presentation Name(s) | RACE |
| Definition | White = a person having origins in any of the original peoples of Europe, the Middle East or North Africa; Black or African-American = A person having origins in any of the black racial groups of Africa; Asian=A person having origins in any of the original peoples of the Far East, Southeast Asia, or the Indian subcontinent including for example, Cambodia, China, India, Japan, Korea, Malaysia, Pakistan, the Philippine Islands, Thailand and Vietnam; American Indian/Alaska Native=A person having origins in any of the original peoples of North and South America (including Central American), and who maintains tribal affiliation or community attachment; Native Hawaiian or Other Pacific Islander=A person having origins in any of the original peoples of Hawaii, Guam, Samoa, or other Pacific Islands. Multiracial = 2 or more of these races (OMB-15, 1997). |
| Valid Values | 1=White, 2=Black or African American, 3=Asian, 4=American Indian/Alaska Native, 5=Native Hawaiian / Pacific Islander, 6=Multiracial; -1=unknown. |
| Data Type | Integer |
| Field Length | 2 |
| Unit of Measurement | Unitless |

| Property | Value |
|---|---|
| Level of Measurement | Nominal |
| Unit of Analysis | Unitless |
| Level of Analysis | Nominal |
| Time Stamp of Standard | 6/12/2000. |

| Property | Value |
|---|---|
| Presentation Name(s) | **ETHNICITY** |
| Definition | Ethnicity, currently limited to asking whether the person is "Hispanic or Latino" (A person of Mexican, Puerto Rican, Cuban, South or Central American, or other Spanish culture or origin, regardless of race) (OMB-15, 1997). |
| Valid Values | 1=Hispanic or Latino, 0=No; -1=unknown. |
| Data Type | Integer |
| Field Length | 2 |
| Unit of Measurement | Unitless |
| Level of Measurement | Nominal |
| Unit of Analysis | Unitless |
| Level of Analysis | Nominal |
| Derivation | Practice in vital records for pre-03 version of certificates: Derived from Origin values 1-5 inclusive. |
| Time Stamp of Standard | 3/16/2000 per U.S. OMB Directive 15, 1997. |

| Property | Value |
|---|---|
| Presentation Name(s) | **SEX** |
| Definition | Biological sex |
| Categorical Attributes | Male / Female |
| Valid Values | 1=Male, 2=Female, -1=unknown. |
| Data Type | Integer |
| Field Length | 2 |
| Unit of Measurement | Unitless |
| Level of Measurement | Nominal |
| Unit of Analysis | Unitless |
| Level of Analysis | Nominal |
| Time Stamp of Standard | 5/4/2000. |

**"Place" requires COLLECTION of 4 data items: Street, City, State, and Zip such that geocoding can assign a latitude/longitude.**
**Specifications for collection of each follow:**

| Property | Value |
|---|---|
| Presentation Name(s) | STREET ADDRESS |
| Definition | Geographic street address of event owner, event, or facility. |
| Valid Values | Street number followed by street name, followed by street direction (e.g. NW), followed by Apartment identifier, followed by room identifier.  Storage of street names will follow standard abbreviations such as RD, BLVD, LN, ST. (See http://www.gis.co.clay.mn.us/usps.htm  ) -1 = Unknown. |
| Data Type | String |
| Field Length | 60 |
| Unit of Measurement | Unitless |
| Level of Measurement | Nominal |
| Unit of Analysis | Unitless |
| Level of Analysis | Nominal |
| Derivation | Self-report. |
| Time Stamp of Standard | 3/16/2000 2.10.2004 – no need to proclaim "0"=Out of state. 6.10.13 – updated spec to be 60 characters instead of 40. |

| Property | Value |
|---|---|
| Presentation Name(s) | CITY NAME |
| Definition | For use in collection of city names: Geographic city or town or parish or Census Designated Place (CDP) or village where the event occurred or event owner lived. |
| Valid Values | USPS City Names; -1 = unknown, 0=non-Georgia resident, _UNIN = resident/event in an unincorporated place (not in a city limit). |
| Data Type | STRING |
| Field Length | 27 |
| Unit of Measurement | Unitless |
| Level of Measurement | Nominal |
| Unit of Analysis | Unitless |
| Level of Analysis | Nominal |
| Derivation | A collection item.  City is Derived from latitude/longitude via a Geocoding process where applicable, and stored as a city code. |
| Time Stamp of Standard | 3/16/2000.  1/23/07 |

| Property | Value |
| --- | --- |
| Presentation Name(s) | **ZIPCODE** |
| Definition | Geographic zipcode+4 where the event owner lived or where event occurred, using USPS standard zipcodes. |
| Valid Values | USPS zipcodes; -1 = unknown.  0=non-Georgia zip. |
| Data Type | String |
| Field Length | 9 |
| Unit of Measurement | Unitless |
| Level of Measurement | Nominal |
| Unit of Analysis | Unitless |
| Level of Analysis | Nominal |
| Derivation | Derived via geocoding where applicable. Originals stored as *_original. |
| Time Stamp of Standard | 3/16/2000 |

| Property | Value |
| --- | --- |
| Presentation Name(s) | **STATE** |
| Definition | Geographic state of residence or event. |
| Valid Values | USPS postal abbreviations. |
| Data Type | String |
| Field Length | 2 |
| Unit of Measurement | Unitless |
| Level of Measurement | Nominal |
| Unit of Analysis | Unitless |
| Level of Analysis | Nominal |
| Derivation | Self-report |
| Time Stamp of Standard | 3/16/2000 |

Place **STORAGE** standards

| Property | Value |
| --- | --- |
| Presentation Name(s) | **LATITUDE** |
| Definition | Latitude of event or residence. |
| Categorical Attributes | N/A |
| Valid Values |  - - -  0 = non-georgia. |
| Data Type | Floating point. |
| Field Length | Float |
| Unit of Measurement | Unitless |
| Level of Measurement | Nominal |
| Unit of Analysis | Unitless |
| Level of Analysis | Nominal |
| Derivation | Derived from geocoding process, using Street, City, State, Zip. |
| Time Stamp of Standard | 3/16/2000.  See issue 177. |
| Access | [  ] Public Use          [X] Datamart Variable |

| Property | Value |
| --- | --- |
| Presentation Name(s) | **LONGITUDE** |
| Definition | Longitude of event or residence. |
| Categorical Attributes | N/A |
| Valid Values | - -  0 = non-Georgia. |
| Data Type | Floating point. |
| Field Length | Float. |
| Unit of Measurement | Unitless |
| Level of Measurement | Nominal |
| Unit of Analysis | Unitless |
| Level of Analysis | Nominal |
| Derivation | Derived from geocoding process, using Street, City, State, Zip. |
| Time Stamp of Standard | 3/16/2000.  See issue 177. |
| Access | [  ] Public Use          [X] Datamart Variable |

| Property | Value |
| --- | --- |
| Presentation Name(s) | **COUNTY** |
| Definition | Geographic county of event or residence. |
| Valid Values | Two digit state FIPS code 00-99 followed by three digit FIPS county code 000-999; range 13001-13321; 0=Non-Georgia county, -1 = unknown. |
| Data Type | String |
| Field Length | 5 |
| Unit of Measurement | Unitless |
| Level of Measurement | Nominal |
| Unit of Analysis | Unitless |
| Level of Analysis | Nominal |
| Derivation | Derived from geocoding where applicable.  Note encapsulated rules for event county in specific instances. |
| Time Stamp of Standard | 3/16/2000.  0 for non-ga 10.21.2003. |

TIME

| Property | Value |
| --- | --- |
| Presentation Name(s) | **DATE** |
| Definition | Date of an event. |
| Valid Values | Months (mm)=01-12 / Days (dd)=01-31 / Year (ccyy)=18yy, 19yy, or 20yy; 12/31/9999=unknown. |
| Data Type | Date |
| Field Length | 10 |
| Unit of Measurement | Day |
| Level of Measurement | Interval |
| Unit of Analysis | Day |
| Level of Analysis | Interval |
| Time Stamp of Standard | 3/27/2002. |