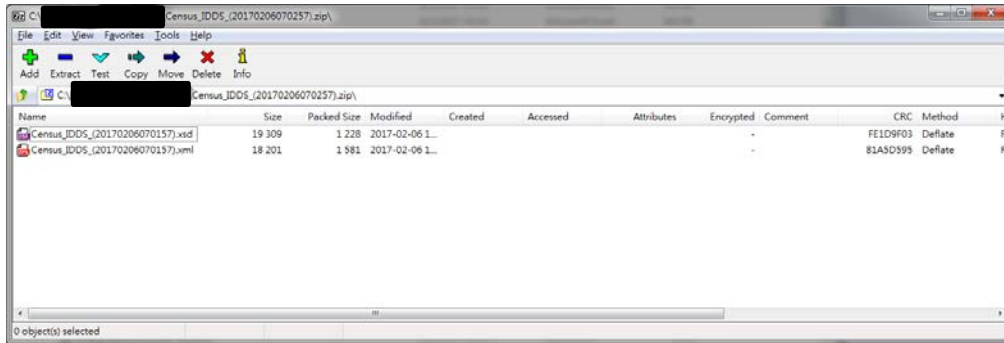


Usage Guide of Statistical Dataset (XML format)

A. Download and Unzip the Statistical Dataset (XML)

1. Download the statistical dataset (in ZIP format) from 2016 Population By-census (16BC) website.
2. Unzip the statistical dataset in a local folder which contains
 - i. 1 XSD file [Census_IDDS_(yyyyMMddhhmmss.xsd)];
 - ii. 1 XML file [Census_IDDS_(yyyyMMddhhmmss.xml)];



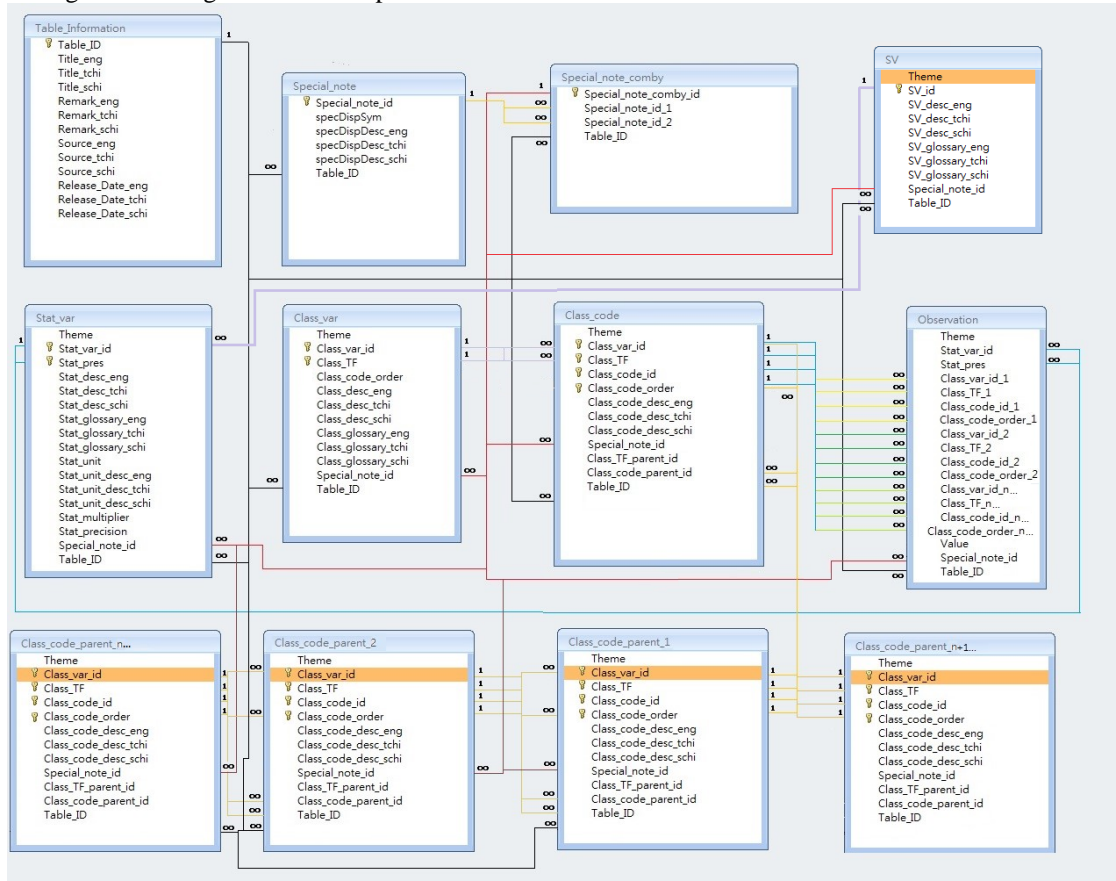
B. XML Schema Definition (XSD)

The specification and relationship of the XML schema are defined as below:

1. XSD element (i.e. data tables)

Class_code
Class_var
Observation
Special_note_comby
Special_note
SV
Stat_var
Table_Information
Class_code_parent_n

2. A diagram showing the relationship between the data tables is shown below:



3. Data Model

The key concepts in the data model are highlighted below.

- 3.1 The aggregated statistics are denoted in the data model as “**Statistical Variables (SV)**”, and further by “**Statistical Presentations (SP)**”. For example, “population” is a SV while “total count”, “total count excluding foreign domestic helpers”, “year-on-year percentage growth” are 3 different SPs of the SV “population”.
- 3.2 Each statistical value, i.e. a special value of a SVSP, is described by one or more “**Classification Variables (CV)**” in order to give a statistical meaning to the value. For example, for the SVSP “population, total count”, a value of “7 071 576” is described (or “classified”) by “year is 2011”; and a value of “3 303 015” is described by “year is 2011” and “sex is male”. Statistical values in more detailed levels are defined by incorporating additional CVs.
- 3.3 A classification variable (CV) can carry different **class code values**, e.g. “male” and “female” are two class code values of the CV “sex”. The class code values of a CV are not necessarily mutually exclusive to each other. For example, for the CV “time”, possible code values are “2001”, “2002”, ..., “2015”; “2015 Jan”, “2015 Feb”, ...; “2015 first half”, “2015 second half”; and for the CV “age”, possible class code values are “0”, “1”, “2”, ..., “84”, “85+”; “0-4”, “5-9”, ..., “80-84”, “85+”; and for the CV “number of household members” (for classifying the SVSP “domestic household, total count”), the possible class code values are “1”, “2”, ..., “5”, “6+”.

- 3.4 A subset of class code values belonging to the same CV can collectively form a **“Tabulation Format (TF)”** of the CV, for the purpose of building a multiple-way statistical table. For example, “population, total count analysed by year (in CCYY format) and quinquennial age group (i.e. “0-4”, “5-9”, ..., “80-84”, “85+”) and sex (“male”, “female”)” is a possible three-way statistical table.
- 3.5 For a single CV, **more than one TF can be defined** to suit different tabulation needs. For example, the SVSP “population, total count” can be analysed by “single-year age” and “sex”; but if “education attainment” is further included as the third CV, the SVSP “population, total count” is only available in “quinquennial age group (i.e. “0-4”, “5-9”, ..., “80-84”, “85+”)”, “sex” and “education attainment”. Hence, two TFs for the CV “age” are required to be defined.
- 3.6 Some TFs (of the same CV) can exhibit **parent-child relationship** between their component class code values. For example, for the CV “age”, the quinquennial age group (i.e. “0-4”, “5-9”, ..., “80-84”, “85+”) is a parent TF to the single-year age (i.e. “0”, “1”, “2”, ..., “84”, “85+”); for the CV “region”, the broad region TF of “Hong Kong, Kowloon, New Territories” is a parent TF to the “District Council District (with 18 districts including, e.g. Central and Western, Wan Chai, Eastern, etc.)”, which in turn is a parent TF to the “Large Tertiary Planning Unit”, and then further to the “Small Tertiary Planning Unit”.

Primary Key (PK) and Foreign Key (FK) Relationship:

Table:	PK:	FK:
Table_Information	Table_ID	
Special_note	Special_note_id	Table ID (Lookup to PK in Table Table_Information)
Special_note_comby	Special_note_comby_id	Table ID (Lookup to PK in Table Table_Information)
		Special_note_id_1 (Lookup to PK in Table Special_note)
		Special_note_id_2 (Lookup to PK in Table Special_note)
SV	SV_id	Table ID (Lookup to PK in Table Table_Information)
		Special_note_id (Lookup to PK in Table Special_note_comby)
Stat_var	Stat_var_id + Stat_pres	Table ID (Lookup to PK in Table Table_Information)
		Stat_var_id (Lookup to PK in Table SV)
		Special_note_id (Lookup to PK in Table Special_note_comby)
Class_var	Class_var_id + Class_TF	Table ID (Lookup to PK in Table Table_Information)
		Special_note_id (Lookup to PK in Table Special_note_comby)
Class_code	Class_var_id + Class_TF + Class_code_id + Class_code_order	Table ID (Lookup to PK in Table Table_Information)
		Special_note_id (Lookup to PK in Table Special_note_comby)
		Class_var_id + Class_TF (Lookup to PK in Class_var)
		Class_var_id + Class_TF_parent_id + Class_code_parent_id (Lookup to PK in Class_code_parent_1)
Class_code_parent_n	Class_var_id + Class_TF + Class_code_id + Class_code_order	Table ID (Lookup to PK in Table Table_Information)
		Special_note_id (Lookup to PK in Table Special_note_comby)
		Class_var_id + Class_TF_parent_id + Class_code_parent_id (Lookup to PK in Class_code_parent_n+1)
Observation		Table ID (Lookup to PK in Table Table_Information)
		Special_note_id (Lookup to PK in Table Special_note_comby)
		Stat_var_id + Stat_pres (Lookup to PK in Table Star_var)
		Class_var_id_1+Class_TF_1+Class_code_id_1 (Lookup to PK in Class_code)
		Class_var_id_2+Class_TF_2+Class_code_id_2 (Lookup to PK in Class_code)
		Class_var_id_n... + Class_TF_n... +Class_code_id_n... (Lookup to PK in Class_code)

C. Table Definition

1. Data Table: Table_Information

This table contains the general information of statistical table.

Table Structure

Fields	Format	PK / FK	Description
Table_ID	Text	PK	Table ID
Title_eng	Text		Title (English)
Title_tchi	Text		Title (Traditional Chinese)
Title_schi	Text		Title (Simplified Chinese)
Remark_eng	Text		Remark (English)
Remark_tchi	Text		Remark (Traditional Chinese)
Remark_schi	Text		Remark (Simplified Chinese)
Source_eng	Text		Source (English)
Source_tchi	Text		Source (Traditional Chinese)
Source_schi	Text		Source (Simplified Chinese)
Release_Date_eng	Text		Release date (English)
Release_Date_tchi	Text		Release date (Traditional Chinese)
Release_Date_schi	Text		Release date (Simplified Chinese)

2. Data Table: Special_note

Special description in statistical table

Table Structure

Fields	Format	PK / FK	Description
Special_note_id	Text	PK	
specDispSym	Text		Notation
specDispDesc_eng	Text		Special description (English)
specDispDesc_tchi	Text		Special description (Traditional Chinese)
specDispDesc_schi	Text		Special description (Simplified Chinese)
Table_ID	Text	FK to Table Table_Information	

3. Data Table: Special_note_comby

This Data Table contains the possible combination of the special note. Individual Special_note_id (e.g. Special_note_id_1 and Special_note_id_2 will be saved as FK in the table Special_note. There are maximum **TWO** special descriptions in one cell.

Table Structure

Fields	Format	PK / FK	Description
Special_note_comby_id	Text	PK	For example, "1,2"
Special_note_id_1	Text	FK to Table Special_note	The first special description key. For example, "1"
Special_note_id_2	Text	FK to Table Special_note	The second special description key. For example, "2"
Table_ID	Text	FK to Table Table_Information	For example, "1,2"

4. Data Table: SV

Description of Statistical Variables in statistical table

Table Structure

Fields	Format	PK / FK	Description
Theme	Text		Theme
SV_id	Text	PK	Statistical Variable ID
SV_desc_eng	Text		Description of statistical variable (English)
SV_desc_tchi	Text		Description of statistical variable (Traditional Chinese)
SV_desc_schi	Text		Description of statistical variable (Simplified Chinese)
SV_glossary_eng	Text		Glossary (English)
SV_glossary_tchi	Text		Glossary (Traditional Chinese)
SV_glossary_schi	Text		Glossary (Simplified Chinese)
Special_note_id	Text	FK to Table Special_note_comby	
Table_ID	Text	FK to Table Table_Information	

5. Data Table: Stat_var

Description of Statistical Presentation in statistical table

Table Structure

Fields	Format	PK / FK	Description
Theme	Text		Theme
Stat_var_id	Text	PK	Statistical variable ID
Stat_pres	Text	FK to Table SV	Statistical presentation ID
Stat_desc_eng	Text		Description of statistical presentation (English)
Stat_desc_tchi	Text		Description of statistical presentation (Traditional Chinese)
Stat_desc_schi	Text		Description of statistical presentation (Simplified Chinese)
Stat_glossary_eng	Text		Glossary (English)
Stat_glossary_tchi	Text		Glossary (Traditional Chinese)
Stat_glossary_schi	Text		Glossary (Simplified Chinese)
Stat_unit	Text		Statistical unit
Stat_unit_desc_eng	Text		Statistical unit description (English)
Stat_unit_desc_tchi	Text		Statistical unit description (Traditional Chinese)
Stat_unit_desc_schi	Text		Statistical unit description (Simplified Chinese)
Stat_multiplier	Numeric		Statistical multiplier
Stat_precision	Numeric		Statistical precision
Special_note_id	Text	FK to Table Special_note_comby	
Table_ID	Text	FK to Table Table_Information	

Statistics Precision

The order of magnitude and the level of rounding of the statistical data are specified by the Multiplier Code (Stat_multiplier) and the Precision Code (Stat_precision) in the Stat_var Table respectively.

Stat_precision	Rounding of figures
-4	Rounded to 4 decimal points
-3	Rounded to 3 decimal points
-2	Rounded to 2 decimal points
-1	Rounded to 1 decimal point
0	Rounded to units place
1	Rounded to tens place
2	Rounded to hundreds place
3	Rounded to thousands place
4	Rounded to ten-thousands place
5	Rounded to hundred-thousands place
6	Rounded to millions place
7	Rounded to ten-millions place
8	Rounded to hundred-millions place
9	Rounded to billions place
10	Rounded to ten-billions place

6. Data Table: Class_var

Description of Classification Variables in statistical table

Table Structure

Fields	Format	PK / FK	Description
Theme	Text		Theme
Class_var_id	Text	PK	Class variable ID
Class_TF	Text		Tabulation format ID
Class_TF_order	Numeric		Tabulation format order
Class_desc_eng	Text		Description of classification variable (English)
Class_desc_tchi	Text		Description of classification variable (Traditional Chinese)
Class_desc_schi	Text		Description of classification variable (Simplified Chinese)
Class_glossary_eng	Text		Glossary (English)
Class_glossary_tchi	Text		Glossary (Traditional Chinese)
Class_glossary_schi	Text		Glossary (Simplified Chinese)
Special_note_id	Text	FK to Table Special_note_comby	
Table_ID	Text	FK to Table Table_Information	

7. Data Table: Class_code

Description of Class Code Value in statistical table

Table Structure

Fields	Format	PK / FK		Description
Theme	Text			Theme
Class_var_id	Text	PK	FK Class_var_id + Class_TF + Class_code_order to Table Class_var	Classification variable ID
			FK Class_var_id + Class_TF_parent_id + Class_code_parent_id to Table Class_code_parents which is the parent of this Class_TF	
Class_TF	Text		FK Class_var_id + Class_TF + Class_code_order to Table Class_var	Tabulation format ID
Class_code_id	Text			Class code ID
Class_code_order	Text			Class code order
Class_code_desc_eng	Text			Description of class code value (English)
Class_code_desc_tchi	Text			Description of class code value (Traditional Chinese)
Class_code_desc_schi	Text			Description of class code value (Simplified Chinese)
Special_note_id	Text	FK to Table Special_note_comby		
Class_var_parent_id	Text	FK Class_var_parent_id + Class_TF_parent_id + Class_code_parent_id of Table Class_code_parents_n which is the parent of this Class_TF		Foreign key of class_code_parents table.
Class_TF_parent_id	Text			
Class_code_parent_id	Text			
Table_ID	Text	FK to Table Table_Information		

8. Data Table: Observation

This data table shows the statistical data in statistical table. Each value is defined by one SVSP and more than one CV or TF.

Table Structure

Fields	Format	PK / FK	Description
Theme	Text		Theme
Stat_var_id	Text	FK to Table Stat_var	Statistical variable ID
Stat_pres	Text		Statistical presentation ID
Class_var_id_1	Text	FK to Table Class_code	Classification variable ID
Class_TF_1	Text		Tabulation format ID
Class_code_id_1	Text		Class code ID
Class_code_order_1	Text		Class code order ID
Class_var_id_2	Text	FK to Table Class_code	If more than one CV or TF
Class_TF_2	Text		
Class_code_id_2	Text		
Class_code_order_2	Text		
Value	Text		- Statistical data in statistical table - Blank if value in statistical table is “ - “
Special_note_id	Text	FK to Table Special_note_comby	
Table_ID	Text	FK to Table Table_Information	

9. Data Table: Class_code_parent_n

This table will be created if there is parent-child relationship between their component class code values

Table Structure

Fields	Format	PK / FK	Description
Theme	Text		Theme
Class_var_id	Text	PK	Classification variable ID
Class_TF	Text		Tabulation format ID
Class_code_id	Text		Class code ID
Class_code_order	Numeric		Class code order
Class_code_desc_eng	Text		Description of class code value (English)
Class_code_desc_tchi	Text		Description of class code value (Traditional Chinese)
Class_code_desc_schi	Text		Description of class code value (Simplified Chinese)
Special_note_id	Text	FK to Table Special_note_comby	
Class_var_parent_id	Text	FK Class_var_parent_id + Class_TF_parent_id + Class_code_parent_id to Table Class_code_parent_n+1	
Class_TF_parent_id	Text		
Class_code_parent_id	Text		
Table_ID	Text	FK to Table Table_Information	