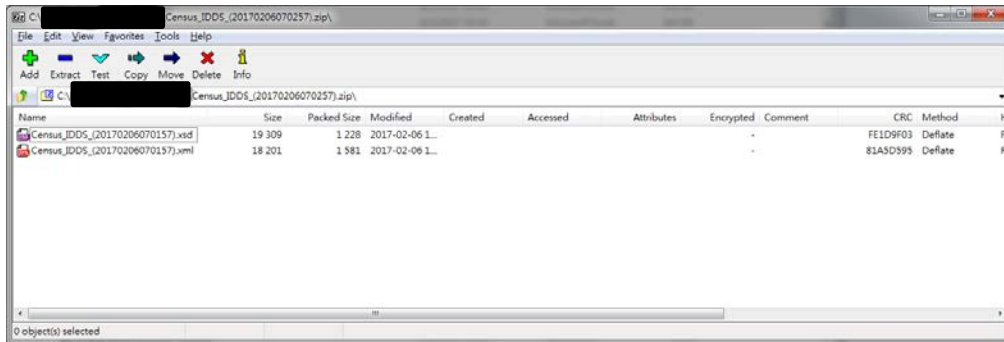


統計數據集(XML 格式)使用指南

A. 下載及解壓統計數據集

1. 從 2016 年中期人口統計的網站下載統計數據集(ZIP 格式)。
2. 把統計數據集解壓到你的電腦內的一個資料夾，其中包括：
 - i. 1 個 XSD 檔案 [Census_IDDS_(yyyyMMddhhmmss.xsd)];
 - ii. 1 個 XML 檔案[Census_IDDS_(yyyyMMddhhmmss.xml)];



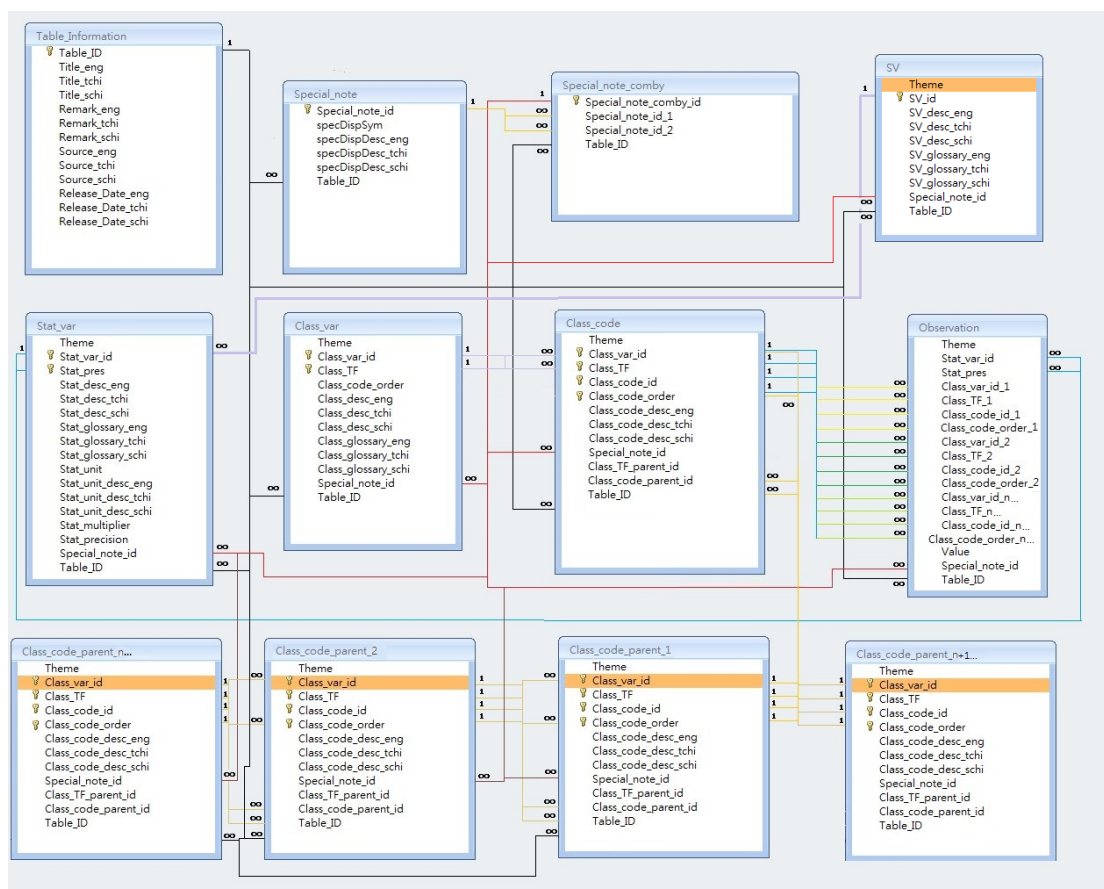
B. XML 結構描述 (XSD)

以下是 XML 結構描述的規格和關係:

1. XSD元素 (即數據表)

Class_code
Class_var
Observation
Special_note_comby
Special_note
SV
Stat_var
Table_Information
Class_code_parent_n

2. 下圖顯示不同數據表之間的關係：



3. 數據模型

以下摘要提供數據模型中的關鍵概念：

- 3.1 在數據模型中，統計數據是由“統計變數(SV)”及“統計展示(SP)”所表示，例如，“人口”是統計變數，而“總數”、“去年同期增長百分比”及“總數(不包括外籍家庭傭工)”是統計變數“人口”的 3 種不同統計展示。
- 3.2 每個統計數值(即統計變數及展示(SVSP)的數值)，是由一個或多個“分類變數(CV)”所定義，以給出該數值的統計意義。例如，對於 SVSP “人口總數”，“7 071 576”的數值是由“年份是 2011”所定義(或“分類”)；而“3 303 015”的數值則由“年份是 2011”和“性別是男性”所定義。統計數值的更詳細級別，可通過包含更多的分類變數而定義。
- 3.3 一個分類變數可以有不同的分類編碼數值(Class code values)，例如，“男性”和“女性”是分類變數“性別”的兩個分類編碼數值。一個分類變數所包含的分類編碼數值不一定彼此互不相容。例如，分類變數“時間”，可能的分類編碼數值為“2001”、“2002”、...、“2015”；“2015 年 1 月”、“2015 年 2 月” ...；“2015 年上半年”、“2015 年下半年”，對於分類變數“年齡”，

可能的分類編碼數值為“0”、“1”、“2”、...、“84”、“85+”;“0-4”、“5-9”、...、“80-84”、“85+”;對於分類變數“家庭成員人數”(用於分類SVSP“家庭住戶總數”),可能的分類編碼數值為“1”、“2”、...、“5”、“6+”。

- 3.4 屬於相同分類變數的分類編碼數值的子集可以共同組成一個“製表格式(TF)”,以用於構建多維統計表。例如,按年份(CCYY格式)、五歲年齡組別(即“0-4”、“5-9”、...、“80-84”、“85+”)和性別(“男”、“女”)分析的人口總數是一個有可能存在的三維統計表。
- 3.5 對於單一個分類變數,可以定義多個製表格式以乎合不同的製表需求。例如,SVSP“人口總數”可以按“單歲年齡組別”和“性別”進行分析;但如果進一步納入“教育程度”作為第三個分類變數,SVSP“人口總數”僅適用於“五歲年齡組別(即“0-4”、“5-9”、...、“80-84”、“85+”)”、“性別”和“教育程度”。因此,分類變數“年齡”需要兩個製表格式作定義。
- 3.6 某些屬於相同分類變數的製表格式可以在其分類編碼數值之間展現主次關係。例如,對於分類變數“年齡”,五歲年齡組別(即“0-4”、“5-9”、...、“80-84”、“85+”)是單歲年齡組別(即“0”、“1”、“2”、...、“84”、“85+”)的主製表格式;對於分類變數“地區”而言,“香港、九龍、新界”是「區議會區分區」(包括中西區,灣仔,東區等.....十八個區)的主製表格式,也是小規劃統計區〔大合併組〕及小規劃統計區〔小合併組〕的主製表格式。

主要鍵 (PK) 和外部鍵 (FK) 關係:

資料表:	主要鍵(PK):	外部鍵(FK):
Table_Information	Table_ID	
Special_note	Special_note_id	Table_ID (搜尋資料表 Table_Information 的主要鍵)
Special_note_comb_y	Special_note_comby_id	Table_ID (搜尋資料表 Table_Information 的主要鍵)
		Special_note_id_1 (搜尋資料表 Special_note 的主要鍵)
		Special_note_id_2 (搜尋資料表 Special_note 的主要鍵)
SV	SV_id	Table_ID (搜尋資料表 Table_Information 的主要鍵)
		Special_note_id (搜尋資料表 Special_note_comby 的主要鍵)
Stat_var	Stat_var_id + Stat_pres	Table_ID (搜尋資料表 Table_Information 的主要鍵)
		Stat_var_id (搜尋資料表 Table SV 的主要鍵)
		Special_note_id (搜尋資料表 Special_note_comby 的主要鍵)
Class_var	Class_var_id + Class_TF	Table_ID (搜尋資料表 Table_Information 的主要鍵)
		Special_note_id (搜尋資料表 Table Special_note_comby 的主要鍵)
Class_code	Class_var_id + Class_TF + Class_code_id + Class_code_order	Table_ID (搜尋資料表 Table_Information 的主要鍵)
		Special_note_id (搜尋資料表 Table Special_note_comby 的主要鍵)
		Class_var_id + Class_TF (搜尋資料表 Class_var 的主要鍵)
		Class_var_id + Class_TF_parent_id + Class_code_parent_id (搜尋資料表 Class_code_parent_1 的主要鍵)
Class_code_parent_n	Class_var_id + Class_TF + Class_code_id + Class_code_order	Table_ID (搜尋資料表 Table_Information 的主要鍵)
		Special_note_id (搜尋資料表 Special_note_comby 的主要鍵)
		Class_var_id + Class_TF_parent_id + Class_code_parent_id (搜尋資料表 Class_code_parent_n+1 的主要鍵)
Observation		Table_ID (搜尋資料表 Table_Information 的主要鍵)
		Special_note_id (搜尋資料表 Special_note_comby 的主要鍵)
		Stat_var_id + Stat_pres (搜尋資料表 Table Stat_var 的主要鍵)
		Class_var_id_1+Class_TF_1+Class_code_id_1 (搜尋資料表 Class_code 的主要鍵)
		Class_var_id_2+Class_TF_2+Class_code_id_2 (搜尋資料表 Class_code 的主要鍵)
		Class_var_id_n... + Class_TF_n... +Class_code_id_n... (搜尋資料表 Class_code 的主要鍵)

C. 資料表定義

1. 資料表: Table_Information

此資料表包含統計表的一般資訊。

資料表結構

資料欄位	格式	主要鍵 (PK) / 外部鍵 (FK)	說明
Table_ID	文字	PK	資料表識別碼
Title_eng	文字		標題 (英文)
Title_tchi	文字		標題 (繁體)
Title_schi	文字		標題 (簡體)
Remark_eng	文字		備註 (英文)
Remark_tchi	文字		備註 (繁體)
Remark_schi	文字		備註 (簡體)
Source_eng	文字		資料來源 (英文)
Source_tchi	文字		資料來源 (繁體)
Source_schi	文字		資料來源 (簡體)
Release_Date_eng	文字		發布日期 (英文)
Release_Date_tchi	文字		發布日期 (繁體)
Release_Date_schi	文字		發布日期 (簡體)

2. 資料表: Special_note

統計表中的特殊說明

資料表結構

資料欄位	格式	主要鍵 (PK) / 外部鍵 (FK)	說明
Special_note_id	文字	PK	
specDispSym	文字		標記
specDispDesc_eng	文字		特殊說明 (英文)
specDispDesc_tchi	文字		特殊說明 (繁體)
specDispDesc_schi	文字		特殊說明 (簡體)
Table_ID	文字	FK (Table_Information)	

3. 資料表: Special_note_comby

此資料表包含特殊注釋的可能性組合。個別的 Special_note_id (例如 Special_note_id_1 和 Special_note_id_2) 將被儲存為資料表 Special_note 的外部鍵。一個儲存格中最多只可有兩個特殊注釋。

資料表結構

資料欄位	格式	主要鍵 (PK) / 外部鍵 (FK)	說明
Special_note_comby_id	文字	PK	例如 "1,2"
Special_note_id_1	文字	FK (Special_note)	第一個特殊注釋的鍵, 例如 "1"
Special_note_id_2	文字	FK (Special_note)	第二個特殊注釋的鍵, 例如 "2"
Table_ID	文字	FK (Table_Information)	例如 "1,2"

4. 資料表: SV

統計表中統計變數的說明

資料表結構

資料欄位	格式	主要鍵 (PK) / 外部鍵 (FK)	說明
Theme	文字		主題
SV_id	文字	PK	統計變數識別碼
SV_desc_eng	文字		統計變數的說明(英文)
SV_desc_tchi	文字		統計變數的說明(繁體)
SV_desc_schi	文字		統計變數的說明(簡體)
SV_glossary_eng	文字		詞彙(英文)
SV_glossary_tchi	文字		詞彙(繁體)
SV_glossary_schi	文字		詞彙(簡體)
Special_note_id	文字	FK(Special_note_comby)	
Table_ID	文字	FK(Table_Information)	

5. 資料表: Stat_var

統計表中統計展示的說明

資料表結構

資料欄位	格式	主要鍵 (PK) / 外部鍵 (FK)		說明
Theme	文字			主題
Stat_var_id	文字	PK	FK (SV)	統計變數識別碼
Stat_pres	文字			統計展示識別碼
Stat_desc_eng	文字			統計展示的說明 (英文)
Stat_desc_tchi	文字			統計展示的說明 (繁體)
Stat_desc_schi	文字			統計展示的說明 (簡體)
Stat_glossary_eng	文字			詞彙 (英文)
Stat_glossary_tchi	文字			詞彙 (繁體)
Stat_glossary_schi	文字			詞彙 (簡體)
Stat_unit	文字			統計單位
Stat_unit_desc_eng	文字			統計單位的說明 (英文)
Stat_unit_desc_tchi	文字			統計單位的說明 (繁體)
Stat_unit_desc_schi	文字			統計單位的說明 (簡體)
Stat_multiplier	數字			數字倍數
Stat_precision	數字			數字精確度
Special_note_id	文字	FK(Special_note_comby)		
Table_ID	文字	FK(Table_Information)		

數字精確度

統計數據的數量級(order of magnitude)和四捨五入層級(level of rounding)分別由資料表 Stat_var 內的倍數編碼(Stat_multiplier)及精確度編碼(Stat_precision)所界定。

Stat_precision	數字的捨入
-4	四捨五入至四個小數點
-3	四捨五入至三個小數點
-2	四捨五入至兩個小數點
-1	四捨五入至一個小數點
0	四捨五入至個位
1	四捨五入至十位
2	四捨五入至百位
3	四捨五入至千位
4	四捨五入至萬位
5	四捨五入至十萬位
6	四捨五入至百萬位
7	四捨五入至千萬位
8	四捨五入至億位
9	四捨五入至十億位
10	四捨五入至百億位

6. 資料表: Class_var

統計表中分類變數的說明

資料表結構

資料欄位	格式	主要鍵 (PK) / 外部鍵 (FK)	說明
Theme	文字		主題
Class_var_id	文字	PK	分類變數識別碼
Class_TF	文字		製表格式識別碼
Class_TF_order	數字		製表格式次序
Class_desc_eng	文字		分類變數的說明 (英文)
Class_desc_tchi	文字		分類變數的說明(繁體)
Class_desc_schi	文字		分類變數的說明(簡體)
Class_glossary_eng	文字		詞彙 (英文)
Class_glossary_tchi	文字		詞彙 (繁體)
Class_glossary_schi	文字		詞彙 (簡體)
Special_note_id	文字	FK(Special_note_comby)	
Table_ID	文字	FK(Table_Information)	

7. 資料表: Class_code

統計表格中分類編碼的說明

資料表結構

資料欄位	格式	主要鍵 (PK) /外部鍵 (FK)	說明
Theme	文字		主題
Class_var_id	文字	PK FK “Class_var_id + Class_TF + Class_code_order” (Class_var 此 Class_TF 的主表)	分類變數識別碼
		FK ‘Class_var_id + Class_TF_parent_id + Class_code_parent_id “ (Class_code_parents)	
Class_TF	文字	FK “Class_var_id + Class_TF + Class_code_order “ (Class_var)	製表格式識別碼
Class_code_id	文字		分類編碼識別碼
Class_code_order	文字		分類編碼次序
Class_code_desc_eng	文字		分類編碼數值的說明 (英文)
Class_code_desc_tchi	文字		分類編碼數值的說明(繁體)
Class_code_desc_schi	文字		分類編碼數值的說明(簡體)
Special_note_id	文字	FK(Special_note_comby)	
Class_var_parent_id	文字	FK “Class_var_id + Class_TF_parent_id + Class_code_parent_id “ (Class_code_parents 此 Class_TF 的主表)	FK(class_code_parents)
Class_TF_parent_id	文字		
Class_code_parent_id	文字		
Table_ID	文字	FK (Table_Information)	

8. 資料表: Observation

此數據表顯示統計表內的統計數據。每個值由一個 SVSP 及多於一個的 CV 或 TF 所定義。

資料表結構

資料欄位	格式	主要鍵 (PK) / 外部鍵 (FK)	說明
Theme	文字		主題
Stat_var_id	文字	FK (Stat_var)	統計變數識別碼
Stat_pres	文字		統計展示識別碼
Class_var_id_1	文字	FK (Class_var)	分類變數識別碼
Class_TF_1	文字		製表格式識別碼
Class_code_id_1	文字		分類編碼識別碼
Class_code_order_1	文字		分類編碼次序
Class_var_id_2	文字	FK (Class_var)	如多於一個 CV 或 TF
Class_TF_2	文字		
Class_code_id_2	文字		
Class_code_order_2	文字		
Value	文字		- 統計表內的統計數據 - 空白如統計表內顯示為“-“
Special_note_id	文字	FK (Special_note_comby)	
Table_ID	文字	FK (Table_Information)	

9. 資料表: Class_code_parent_n

如在分類編碼數值間存在主次關係，此資料表將會被建立。

資料表結構

資料欄位	格式	主要鍵 (PK) / 外部鍵 (FK)	說明
Theme	文字		主題
Class_var_id		PK	分類變數識別碼
Class_TF	文字		製表格式識別碼
Class_code_id	文字		分類編碼識別碼
Class_code_order	數字		分類編碼次序
Class_code_desc_eng	文字		分類編碼數值的說明(英文)
Class_code_desc_tchi	文字		分類編碼數值的說明(繁體)
Class_code_desc_schi	文字		分類編碼數值的說明(簡體)
Special_note_id	文字	FK (Special_note_comby)	
Class_var_parent_id	文字	FK “Class_var_parent_id + Class_TF_parent_id + Class_code_parent_id” (Class_code_parent_n+1)	
Class_TF_parent_id	文字		
Class_code_parent_id	文字		
Table_ID	文字	FK (Table_Information)	