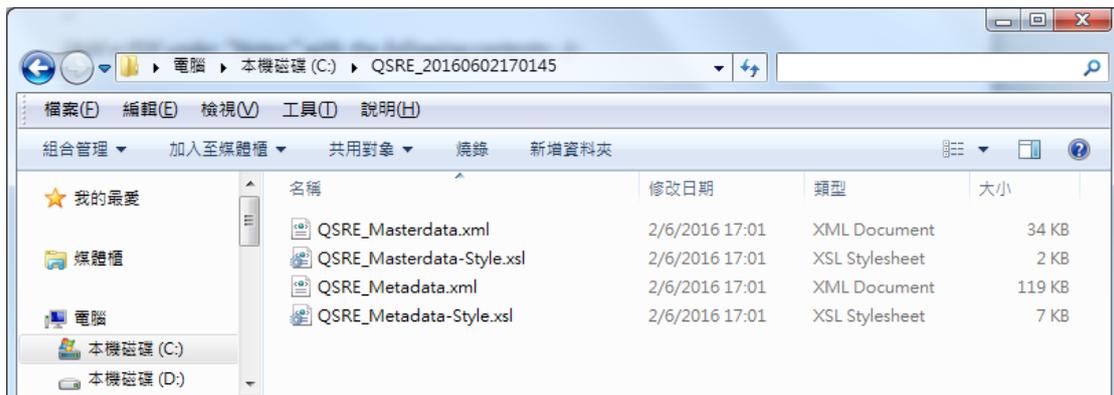


統計數據集(XML 格式)使用指南

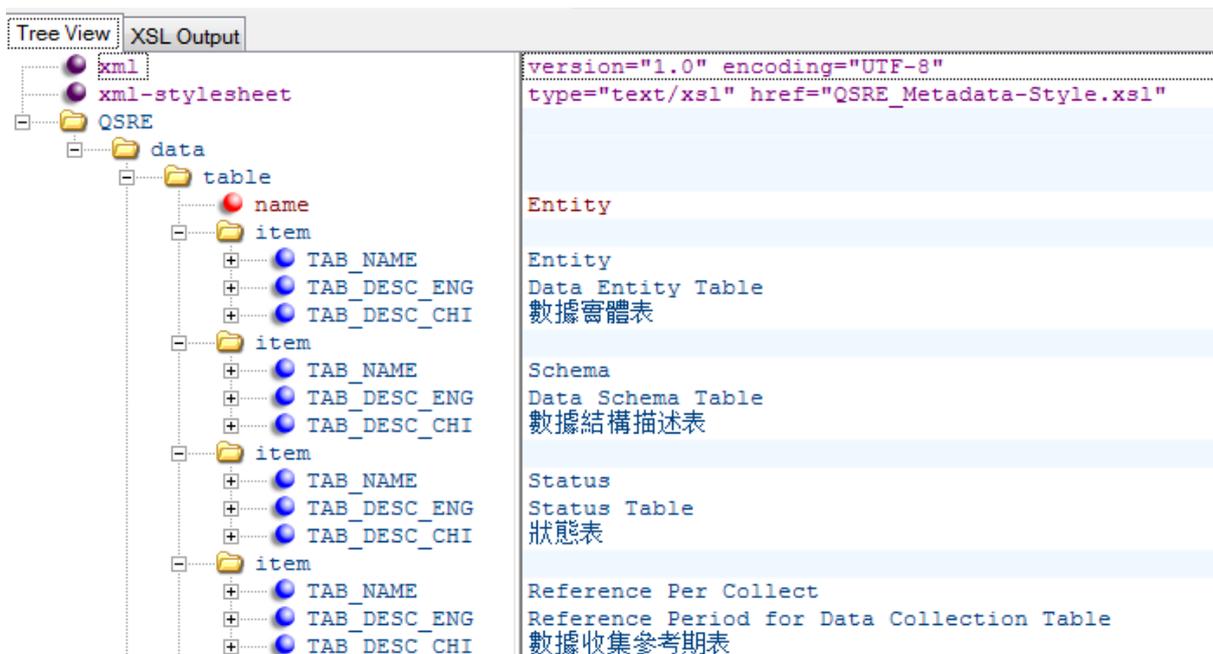
A. 下載及解壓統計數據集

1. 從政府統計處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：
 - 1 個元數據集(Metadata dataset) (XML 格式)
 - 1 個或多個主要數據集(Master Data dataset(s)) (XML 格式)
 - 1 個元數據集的樣式表(Stylesheet File) (XSL 格式)
 - 1 個主要數據集的樣式表(Stylesheet File) (XSL 格式)



B. 元數據集(Metadata Dataset)

3. 元數據集是統計數據集的數據字典，提供主要數據集(Master Data dataset)內的統計數據的規格和描述。



```
version="1.0" encoding="UTF-8"
type="text/xsl" href="QSRE_Metadadata-Style.xsl"

Entity
Entity
Data Entity Table
數據實體表

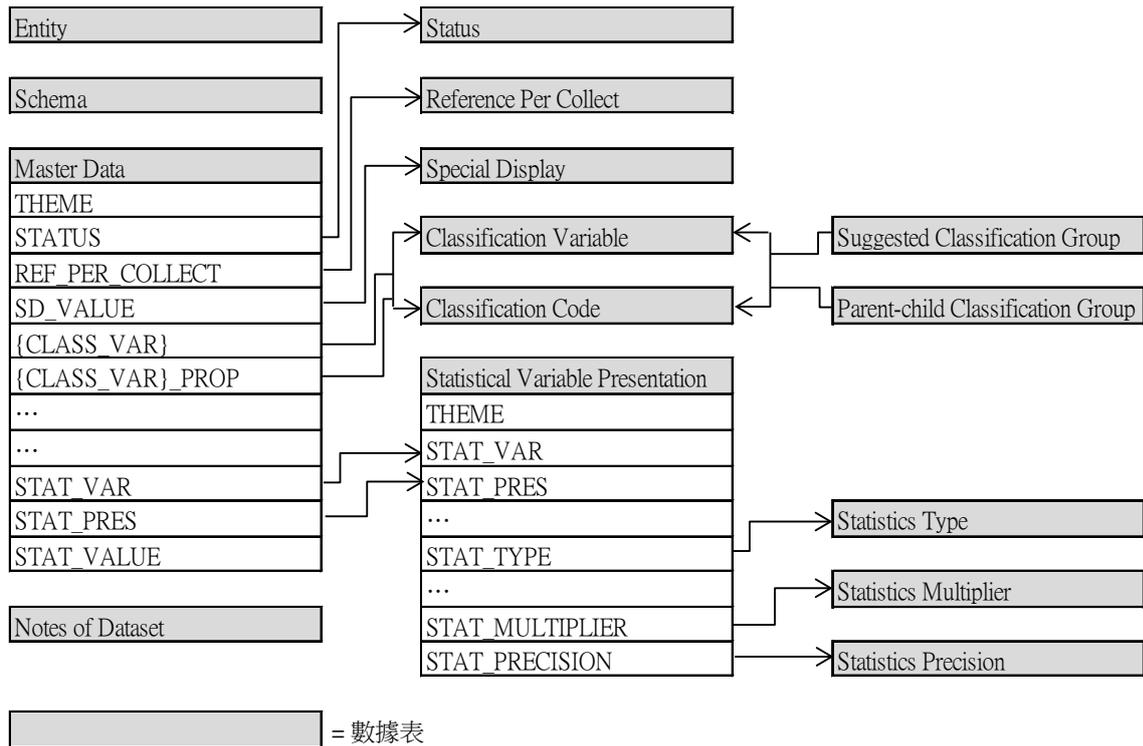
Schema
Data Schema Table
數據結構描述表

Status
Status Table
狀態表

Reference Per Collect
Reference Period for Data Collection Table
數據收集參考期表
```

(使用 XML 編輯軟件(如 Microsoft XML Notepad)開啟元數據集的螢幕畫面)

4. 下圖顯示不同數據表之間的關係:



5. 實體表(Entity table)和結構描述表(Schema table)詳列每個數據表(包括主要數據集)的表名稱、欄名稱及欄種類。此外，數據集註釋表(Notes of Dataset table)提供關於該統計數據集需要注意的要點(如有)。

C. 主要數據集(Master Data Dataset)

6. 主要數據集提供個別統計數據的數值。

(使用 XML 編輯軟件(如 Microsoft XML Notepad)開啟主要數據集的螢幕畫面)

7. 每個統計數值(STAT_VALUE)的屬性由以下項目具體說明:
- 統計變數(STAT_VAR)，以及其特定的統計變數展示編碼(STAT_PRES)如原始數據、比例數據或變化率，這些變數及編碼是由統計變數展示表(Statistical Variable Presentation table)和統計數字類型表(Statistics Type table)所定義；
 - 分類變數{CLASS_VAR}及其相應的分類編碼{CLASS_CODE}，包括統計數據的參考時間，這些變數及編碼是由分類變數表(Classification Variable table)和分類編碼表(Classification Code table)所定義；
 - 特殊顯示編碼(SD_VALUE)說明統計數據的特別備註(如有)，這些編碼是由特殊顯示表(Special Display table)所定義；及
 - 其他編碼如狀態編碼(STATUS)及數據收集參考期編碼(REF_PER_COLLECT)，這些編碼分別由狀態表(Status table)和數據收集參考期表(Reference Per Collect table)所定義。
8. 另外，統計數據的數量級(order of magnitude)和四捨五入層級(level of rounding)分別由統計變數展示表(Statistical Variable Presentation table)內的倍數編碼(STAT_MULTIPLIER)及精確度編碼(STAT_PRECISION)所界定，這些編碼分別由統計數字倍數表(Statistics Multiplier table)和統計數字精確度表(Statistics Precision table)所定義。
9. 就分類變數和分類編碼而言，建議分類組別表(Suggested Classification Group table)及主次分類編碼組別表(Parent-child Classification Group table)分別提供常用分類編碼組別及分類編碼的主次關係作參考。
10. 使用者應注意統計數據的屬性，而將數據作進一步運算前應多加謹慎。

D. 樣式表(Stylesheet Files)

11. 為方便一般使用者閱讀數據集，樣式表(XSL 格式)可將 XML 格式數據集在 Excel、XML 編輯軟件或網頁瀏覽器(20MB 以下的 XML 數據集適用)內以結構化模式顯示。一般情況下，當 XML 格式數據集與相應的樣式表(XSL 格式)放置在相同的資料夾內，使用者使用瀏覽器或 XML 編輯軟件開啟數據集時，數據集便會自動以列表形式顯示。

The screenshot shows a web browser window with two tables. The first table, titled 'Entity', lists various data tables with their English and Chinese descriptions. The second table, titled 'Schema', lists columns for the 'Entity' table with their descriptions, data types, and primary key status.

TAB_NAME	TAB_DESC_ENG	TAB_DESC_CHI
Entity	Data Entity Table	數據實體表
Schema	Data Schema Table	數據結構描述表
Status	Status Table	狀態表
Reference Per Collect	Reference Period for Data Collection Table	數據收集參考期表
Special Display	Special Display Table	特殊顯示表
Statistics Multiplier	Statistics Multiplier Table	統計數字倍數表
Statistics Precision	Statistics Precision Table	統計數字精確度表
Statistics Type	Statistics Type Table	統計數字類型表
Statistical Variable Presentation	Statistical Variable Presentation Table	統計變數展示表
Classification Variable	Classification Variable Table	分類變數表
Classification Code	Classification Code Table	分類編碼表
Suggested Classification Group	Suggested Classification Group Table	建議分類組別表
Parent-child Classification Group	Parent-child Classification Group Table	主次分類編碼組別表
Notes of Dataset	Notes of Dataset Table	數據集註釋表
Master Data	Master Statistics Table	主要統計數字表

TAB_NAME	COL_NAME	COL_DESC_ENG	COL_DESC_CHI	COL_TYPE	PK
Entity	TAB_NAME	Table name	資料表名稱	Varchar(30)	Y
Entity	TAB_DESC_ENG	Table description in English	資料表的英文描述	Varchar(250)	N
Entity	TAB_DESC_CHI	Table description in Chinese	資料表的中文描述	NVarchar(250)	N

(使用瀏覽器開啟元數據集的螢幕畫面)

E. 主要數據集內的比例類別統計數據

12. 在定義比例類別(STAT_PRES = PROP_{xxx})統計數據(如百分比)的屬性時，是需要額外指明該數據的分母的覆蓋範圍。在主要數據集的每個分類變數{CLASS_VAR}，會附有對應的另一欄，名為“{CLASS_VAR}_PROP”作為分母指標，顯示比例類別統計數據中分母所覆蓋範圍。
13. 如某分類編碼出現於分母指標({CLASS_VAR}_PROP)，意指該分類編碼用作比例類別數據的分母；否則，當分母指標顯示為空白值，分母便覆蓋有關分類變數的所有編碼。
14. 附件一提供一例子作說明之用。

(2018.12)

主要數據集內比例類別統計數據的例子

下表顯示按職業和性別劃分的假定就業人數，供說明之用：

按職業和性別劃分的就業人數

職業(OCC)	性別(SEX)		總數
	男(M)	女(F)	
經理(Manager)	22 (a)	28 (b)	50 (c)
專業人員(Professional)	65 (d)	35 (e)	100 (f)
...
...
總數	400 (g)	600 (h)	1000 (i)

在主要數據集內，按職業(OCC)和性別(SEX)分類變數劃分的就業人數(STAT_VAR= EM(Employed Persons))所佔比例(STAT_PRES=Prop_1dp)會以兩欄(OCC_PROP 和 SEX_PROP)作為分母指標。下表說明各就業人數所佔比例的數值(STAT_VALUE)的屬性。

主要數據集內的比例類別統計數據											
					分類變數及其相應分母指標						
THEME	STATUS	...	CCYY	OCC	OCC_PROP	SEX	SEX_PROP	STAT_VAR	STAT_PRES	STAT_VALUE
XXX	f	...	2015	...	1 (Manager)		F (Female)	F (Female)	EM (Employed Persons)	Prop_1dp	4.7
XXX	f	...	2015	...	1 (Manager)		M (Male)	M (Male)	EM (Employed Persons)	Prop_1dp	5.5
XXX	f	...	2015	...	2 (Professional)	2 (Professional)	M (Male)		EM (Employed Persons)	Prop_1dp	65
XXX	f	...	2015	...	2 (Professional)	2 (Professional)	F (Female)		EM (Employed Persons)	Prop_1dp	35
XXX	f	...	2015	...	1 (Manager)		F (Female)		EM (Employed Persons)	Prop_1dp	2.8

比例的數值(STAT_VALUE)是根據以下公式計算

"4.7"=女性經理在所有職業的女性的百分比
 $(=b)/(h) \times 100\%$
 即分母是所有職業 (OCC_PROP="") 的就業女性 (SEX_PROP="F")

"5.5"=男性經理在所有職業的男性的百分比
 $(=a)/(g) \times 100\%$
 即分母是所有職業 (OCC_PROP="") 的就業男性 (SEX_PROP="M")

"65"=男性專業人員佔所有性別的專業人員的百分比
 $(=d)/(f) \times 100\%$
 即分母是所有性別 (SEX_PROP="") 的專業人員 (OCC_PROP="2")

"35"=女性專業人員佔所有性別的專業人員的百分比
 $(=e)/(f) \times 100\%$
 即分母是所有性別 (SEX_PROP="") 的專業人員 (OCC_PROP="2")

"2.8"=女性經理佔所有就業人士(所有職業及性別)的百分比
 $(=b)/(i) \times 100\%$
 即分母是所有職業 (OCC_PROP="") 及所有性別 (SEX_PROP=" ") 的就業人士