

创建项目数据网页

一、项目数据网页简介：

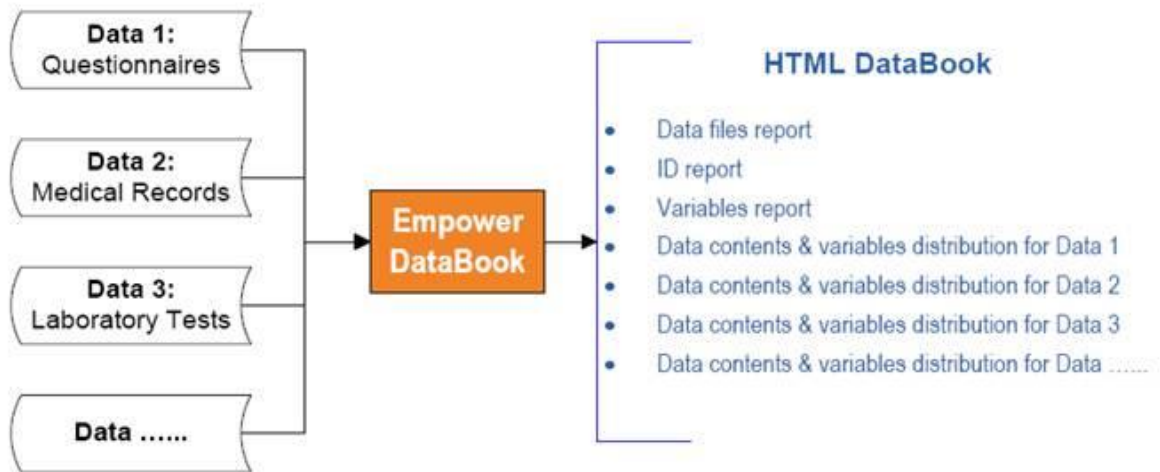
通常一个大型流行病学研究的数据来源于不同地方，如问卷数据、实验室检测数据、随访调查数据、病历摘录数据等等，每个数据存放在不同的数据文件中，这些数据文件通过研究对象编码可以整合到一起。

在数据采集、录入、数据分析过程中，我们常常需要了解：

- (1) 不同源的数据各自进展与样本量。
- (2) 各数据所包含的变量及每个变量的数据分布情况。
- (3) 项目整体数据完整情况，整合多个数据文件时可用的完整数据样本量。

项目数据网页正是为提供查阅上述信息而设计的，它由一系列相互链接的HTML文件组成。这些HTML文件包括：

- (1) 每个数据文件有一个数据内容查阅文件。含样本量、变量名与注解、各变量的分布。
- (2) 整体报告文件。含数据文件列表、研究对象编码列表、研究对象各数据文件完整情况列表。
- (3) 变量名列表文件。含各数据文件所包含的所有变量名及其在各文件中的重叠情况。



二、创建方法：

1. 文件整理：

首先创建一个文件夹（文件夹名最好与研究方案名相同），将该研究项目的所有文件存放到同一文件夹内。文件包含：

- (1) 数据文件（必有）：每个数据文件均应存成制表符分隔的文本文件。
- (2) 变量说明文件（可选）：每个数据文件对应一个变量说明文件，变量说明文件取名为数据文件名后面加_variables.xls。如果数据文件名为：212.xls, 变量说明文件名则应为：212_variables.xls。变量说明文件也应存成制表符分隔的文本文件，格式如下图所示：

变量名	取值编码	意义
FMYTYPE		FMYTYPE
	0	0
	1	1
FMYID		FMYID
SUBJ		SUBJ
NID		NID
	1	1
	2	2
	11	11
	12	12
	13	13
	14	14
	15	15
	16	16
	17	17
SEX		性别
	1	男
	2	女

第一行为列标题，含3列，第一列为变量名，第二列为取值编码，第三列为注解。自第二行开始为每个变量及其编码的注解。每个连续性变量只占一行，其第二列为空，第三列是变量注解；每个分类型变量占多行，首先是该变量的注解，变量注解行的第2列为空，紧接着该变量的每个取值编码的注解，每个编码占一行，编码注解行的第一列为空。

- (3) 问卷/数据采集表单（可选）：每个数据文件对应一个表单文件，存成pdf文件，文件名应与数据文件完全相同，如数据文件名为：212.xls, 表单文件则应为212.pdf。
- (4) 项目方案文件（可选）：将该项目的实施方案存成pdf文件，取名为研究项目名加_protocol.pdf，如项目名为：test，则研究方案文件名应为：test_protocol.pdf。

2. 创建数据网页模块应用

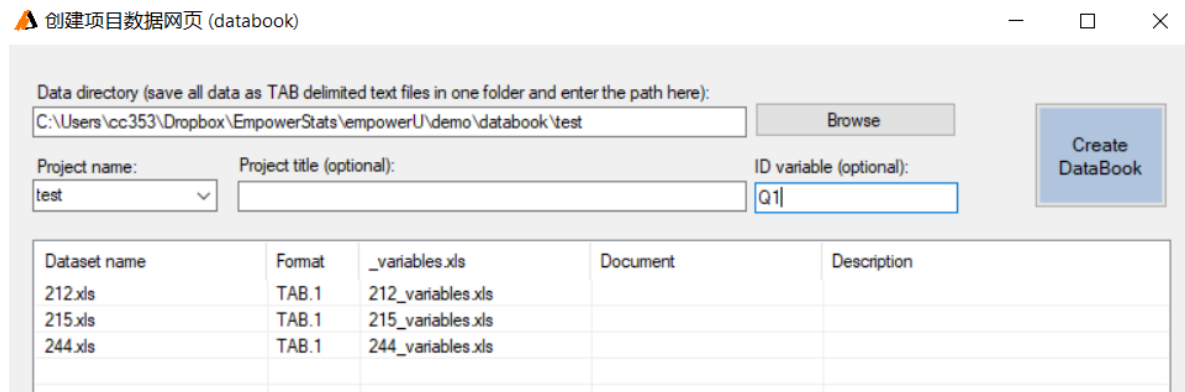
打开易佻统计软件—工具菜单—创建数据网页模块，选择项目文件夹。项目名自动填写为文件夹名，文件夹中的数据文件与变量说明文件自动被查找并填写到数据文件列表中。右击数据文件列表，可以对所选的数据文件进行文字描述，也可剔

除该文件。研究对象编码变量名（可选）须手动填写，如果没有不填研究对象编码将无法对整体数据完整情况进行统计。

例：练习数据 test 项目含 3 个数据文件，每个文件中的变量 Q1 表示研究对象编号。下载地址：

http://www.empowerstats.com/empowerStats/exdata/databook_test.zip

下载后解压到一新子目录中。点击工具菜单，打开创建数据网页。给出该目录地址，如下图：



Dataset name	Format	_variables.xls	Document	Description
212.xls	TAB.1	212_variables.xls		
215.xls	TAB.1	215_variables.xls		
244.xls	TAB.1	244_variables.xls		

数据文件格式是软件自动检测的。变量说明文件也是由软件自动查找的。

点击“Create DataBook”即可。数据网页生成后存放在 C 盘我的文件夹下的 \EmpowerStats\databook\项目名目录下，如上例，项目名为 test，数据网页路径则为：我的文件夹\EmpowerStats\databook\test\test.htm。

点击[这里](#)查看示例项目 test 的数据网页页面：

<http://www.empowerstats.com/cn/manuals/databook/test/test.htm>