

## الفصل السابع

### معالجة البيانات

إن نظام معالجة بيانات الدورة الثالثة من المسوحات العنقودية متعددة المؤشرات مصمّم بحيث يتم تسليم النتائج الأولية لمسح ما ضمن أسابيع قليلة من نهاية العمل الميداني. يتضمن هذا الفصل معلومات سوف تساعدك على أن تتولى التخطيط ودفع عملية التحضير قدماً لتحقيق الهدف. يبدأ الفصل بإعطائك ملخصاً عن نظام معالجة بيانات الدورة الثالثة من المسوحات العنقودية متعددة المؤشرات، ثم يناقش كل جزء من الأجزاء المكونة لها بالتفصيل، مقدّماً لك مراجع لمصادر تكميلية للمعلومات حيثما كان ذلك ملائماً. ينتهي الفصل بمجموعة من ثلاثة قوائم تحقق سوف تساعدك على جعل معالجة البيانات لمسحك عملية ناجحة.

### نظرة عامة

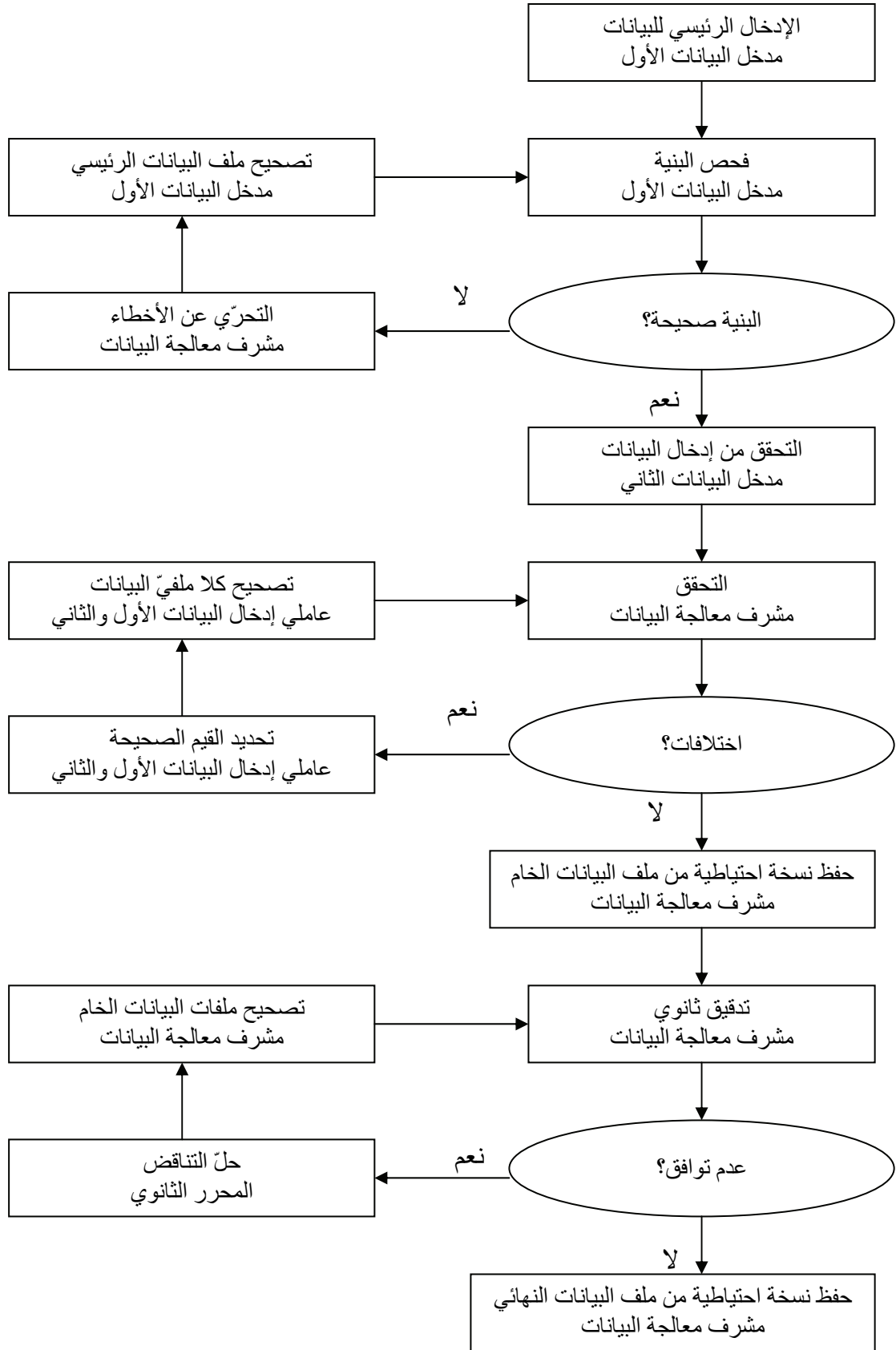
إن السبب في إمكانية أن يحرز نظام معالجة بيانات الدورة الثالثة من المسوحات العنقودية متعددة المؤشرات تتابعاً سريعاً مع الوقت هو أن عملية معالجة البيانات تتم بالتوازي مع العمل الميداني للمسح. لإنجاز ذلك، يتم تخزين بيانات كل عنقود في ملف مستقل. تقوم هذه الطريقة بتجزئة معالجة البيانات إلى مقاطع منفصلة، وتسمح لعملية معالجة البيانات بالتقدّم في الوقت الذي يكون العمل الميداني ما يزال مستمراً. وبالتحديد، تتم عملية معالجة بيانات كل عنقود حالماً تُعاد الاستبيانات من الميدان. وهكذا، فإنه بحلول موعد انتهاء الاستبيانات الأخيرة وإعادتها إلى المراكز الرئيسية، تكون معظم البيانات قد تمت معالجتها بالفعل.

إن عملية معالجة البيانات تبعاً للعنقود ليست عملية صعبة، إلا أنها تحتاج إلى تنظيم شديد الدقة. يمكن تقسيم نظام معالجة البيانات إلى ثلاثة أطوار: طور التحضير وطور المعالجة الأولية للبيانات وطور المعالجة الثانوية للبيانات. وقد قمنا بتلخيص كل من هذه الأطوار في أحد الأقسام الفرعية المذكورة أدناه، ولكل منها قائمة تحقق مرتبطة بها في نهاية هذا الفصل.

### التحضير لإدخال البيانات

إن الهدف من طور التحضير لإدخال البيانات هو الاستعداد للبدء في ادخال البيانات بعد فترة قصيرة من بدء العمل الميداني. يشمل طور التحضير الخطوات التالية:

- الحصول على أجهزة حواسيب وإعداد غرفة لمعالجة البيانات.
- تعيين وتوظيف الأفراد المناسبين للعمل.
- تعديل البرامج حسب الاستبيانات الخاصة بالبلد المنفذ.
- إعداد نظام لإدارة الاستبيانات وملفات البيانات.



## المعالجة الأولية للبيانات

الهدف من المعالجة الأولية للبيانات هو إنتاج ملفات بيانات نظيفة ومدققة. وتتضمن عملية المعالجة الأولية للبيانات الخطوات التالية:

- إدخال جميع الاستبيانات لعنقود ما في ملف بيانات.
- فحص بنية ملف البيانات.
- إدخال البيانات مرة ثانية ثم التحقق من ملف البيانات ( مقارنة الملفين و حل الاختلافات).
- حفظ نسخة احتياطية من ملف البيانات المتحقق منه.
- إجراء عملية التدقيق الثانوي على ملف البيانات.
- حفظ نسخة احتياطية من ملف البيانات المدقق و النهائي.

لقد قمنا بتلخيص تسلسل خطوات عملية معالجة البيانات في مخطط التدفق على الصفحة السابقة. لاحظ بدقة بأن فحص البنية والتحقق من إدخال البيانات والتدقيق الثانوي هي إجراءات قابلة للإعادة يتم تكرارها حتى الوصول إلى حلّ جميع المشاكل أو حتى يتم الإقرار بأن جميع البيانات قد أصبحت مقبولة و متوافقة.

## المعالجة الثانوية للبيانات

إن الهدف من المعالجة الثانوية للبيانات هو إنتاج ملفات بيانات قابلة للتحليل وإنشاء جداول نموذجية للدورة الثالثة من المسوحات العنقودية متعددة المؤشرات. تتضمن المعالجة الثانوية للبيانات الخطوات التالية:

- وضع جميع ملفات بيانات العنقود متسلسلة في ملف بيانات واحد.
- تحويل البيانات إلى حزمة التحليل الاحصائي.
- حساب أوزان العينات.
- حساب معامل الثروة.
- إعادة ترميز البيانات لتسهيل تحليل البيانات.
- إنشاء الجداول اللازمة لتحليل البيانات.

- أرشفة وتوزيع ملفات البيانات.

## الأفراد والبنية التحتية

### الأفراد

يتضمن فريق معالجة البيانات للدورة الثالثة من المسوحات العنقودية متعددة المؤشرات أربعة أنواع من الأفراد: منظم الاستبيانات و مدخلو البيانات ومدققون ثانويون ومشرف معالجة البيانات. وتقع على كل موقع من هذه المواقع الوظيفية مسؤوليات مميزة، ومن المحتمل أن يضرّ الجمع بينها بجودة البيانات.

يقوم **منظم الاستبيانات (أو مدقق مكتبي)** بعمليات الفحص وتنظيم الاستبيانات حينما تصل من الميدان. فعندما يصل عنقود ما إلى مكتب معالجة البيانات، يقوم منظم الاستبيانات بالتحقق من أن جميع الاستبيانات موجودة وجاهزة للإدخال. فإذا وجدت أية استبيانات مفقودة، يجب على المنظم حلّ المشكلة بمساعدة فريق العمل الميداني (الخطوات المحددة التي يجب على منظم الاستبيانات اتخاذها موضحة فيما بعد).

ويقوم **مدخلو البيانات** بإدخال البيانات، ويجب أن تكون لديهم خبرة سابقة في إدخال البيانات وأن يكونوا على معرفة واطلاع بالاستبيانات. أما إحدى الوسائل لإحراز المتطلب الثاني فهو جعل عاملي إدخال البيانات يحضرون دورة تدريب الباحثات. وقبل البدء بإدخال البيانات، يجب إقامة دورة تدريبية منفصلة مدتها يومين أو ثلاثة أيام لمدخلي البيانات لتمكينهم من التعرف على برنامج إدخال البيانات و نظام معالجة البيانات. وفي نهاية التدريب، يجب أن يكون عاملي إدخال البيانات منسجمين بالعمل على برنامج إدخال البيانات ومدركين لمسؤولياتهم اليومية. يعتمد عدد مدخلو البيانات المطلوبين لهذا العمل على عدد أجهزة الحاسوب المتوفرة. وقد تمت مناقشة هذا الموضوع بالتفصيل أدناه.

أما **المدققون الثانويون** فيقومون بالتحريّ عن مشاكل عدم التوافق (التناقضات) المعقدة التي يكشفها برنامج التدقيق الثانوي ويعملون على حلّها. ويجب أن تكون لدى المدققين الثانويين فهماً كاملاً للاستبيانات ولأهداف المسح. وقد قمنا بتقديم إرشادات للتدقيق الثانوي في الملحق السابع وذلك لمساعدة المدققين الثانويين أثناء عملية التدقيق الثانوي. يحتاج المسح النموذجي إلى مدقق ثانوي واحد أو اثنان.

وأما **المشرف على معالجة البيانات**، فهو عضو ذو أهمية بالغة من أعضاء فريق معالجة البيانات، فهو الذي يقوم بتعديل البرامج النموذجية لكي تتلاءم مع الاستبيانات الخاصة بالبلد ومراقبة جميع مهام عمليات معالجة البيانات. يجب أن تكون لدى المشرف على معالجة البيانات خبرة في إدارة معالجة البيانات لمسح كبير أو لتعداد سكاني وفهماً تاماً للاستبيانات بالإضافة إلى مهارات البرمجة في حزمة البرامج CSPro و حزمة التحليل الاحصائي SPSS. ويجب أن يكون المشرف على معالجة البيانات متواجد بصفة دائمة خلال الفترة التي يتم فيها إدخال البيانات وتدقيقها وجدولتها.

يجب تعيين المشرف على معالجة البيانات في مرحلة مبكرة من مراحل التخطيط للمسح حتى يكون على صلة تامة بعملية تنقيح استبيانات الدورة الثالثة من المسوحات العنقودية متعددة

المؤشرات. ويجب استشارة هذا الشخص لضمان أن الرموز التي قد تم استخدامها في الاستبيانات متسقة وليس فيها لبس أو غموض وبأنه قد تم وضع البيانات التعريفية اللازمة جميعها على المشرف على معالجة البيانات أيضاً أن يكون قادراً على المساعدة في عملية المراجعة النهائية للاستبيانات بناء على الخبرة التي اكتسبها أثناء إدخال الاستبيانات في مرحلة الاختبارات القبلية.

## أجهزة الحاسوب والأجهزة الصلبة الأخرى

تبيّن القائمة أدناه الأجهزة اللازمة لمعالجة البيانات:

- أجهزة حاسوب لإدخال البيانات.
- جهاز حاسوب المشرف على معالجة البيانات.
- أداة تخزين ثانوية (مثلاً منفذ الناقل التسلسلي العام (USB) والقرص المدمج القابل لإعادة الكتابة والتخزين (CD-RW).
- أقراص (أو وسيلة تتوفر لمُدخلي البيانات لنقل الملفات إلى جهاز الحاسوب الخاص بالمشرف على معالجة البيانات، عن طريق اتصال الأجهزة بشبكة واحدة على سبيل المثال).
- طابعة.
- ورق.
- وحدة حبر، ريبونات حبر طابعة.
- واقيات من شدة التيار.
- موصل للكهرباء غير منقطع الطاقة (UPS).
- أقلام حبر أخضر.

يجب أن تحتوي حواسيب إدخال البيانات على معالجات بنتيوم Pentium، و نظام تشغيل ويندوز 95 أو أعلى، ذاكرة وصول عشوائي RAM سعة 32 ميغابايت أو أكثر، ومساحة حرّة في القرص الصلب سعة 1 غيغابايت أو أكثر وإما أداة تشغيل قرص لين 3.5" (أو أن تكون الحواسيب متصلة بشبكة واحدة). أما عدد أجهزة حاسوب إدخال البيانات اللازمة لمعالجة المسح فتعتمد على حجم العينة وعلى عدد الساعات التي يعمل فيها مدخل البيانات أسبوعياً والمساحة المتوفرة والجدول الزمني للمسح. للحصول على الرقم التقديري لعدد أجهزة الحاسوب اللازمة لإدخال البيانات، عليك تقدير المدة التي سوف يحتاجها مدخل البيانات لإدخال الاستبيانات النموذجية لأسرة واحدة ثم ضرب هذا الرقم في عدد الأسر المتوقع بموجب تصميم العينة. فإذا لم تكن قادراً على تقدير الوقت الذي يستغرقه إدخال الاستبيانات، قم باستخدام 20-30 دقيقة لكل أسرة كدليل تقريبي اعتماداً على عدد النساء والأطفال المتوقع وجودهم في كل أسرة. في هذه الحالة يتم تقدير عملية الضرب بعدد الأسر للحصول على إجمالي عدد الساعات اللازمة لإدخال البيانات. قم بتقسيم هذا الرقم التقديري بعدد الساعات التي سوف يعمل فيها كل مدخل بيانات في

الأسبوع ثم القسمة على عدد الأسابيع اللازمة لإتمام إدخال البيانات إذا كنت تعترض إتمام إدخال البيانات بعد أسبوع واحد فقط من إعادة آخر استبيان من الميدان. على سبيل المثال، إذا كان حجم العينة للمسح هو 6000 أسرة وتستغرق كل أسرة 20 دقيقة لإدخال بياناتها، يكون الوقت الإجمالي اللازم لإدخال جميع الأسر هو 2000 ساعة. فإذا كان لديك 8 أسابيع لإتمام عملية إدخال البيانات، عندها يلزمك 250 ساعة كل أسبوع. إذا كان مدخل البيانات يعمل 40 ساعة في الأسبوع، فسوف تحتاج إلى 7 أجهزة حاسوب و 7 مدخلي بيانات. أحياناً من الممكن تنظيم فترتي مناوبة لإدخال البيانات حتى يتم استخدام حاسوب واحد بواسطة اثنين من مدخلي البيانات يوميا. يعمل كل مدخل بيانات لمدة، 6 ساعات فيكون الحاسوب قد تم استخدامه مدة 12 ساعة يوميا.

أما حاسوب المشرف فيجب أن يمتلك معالج بيانات أسرع، و نظم تشغيل ويندوز 98 فأعلى، وذاكرة وصول عشوائي سعة 64 ميغابايت أو أكثر ومساحة حرّة في القرص الصلب سعة 1 غيغابايت فأكثر وأداة تشغيل قرص لين 3.5" ( أو شبكة متصلة مع حواسيب إدخال البيانات) وأداة تخزين ثانوية.

وجود موصل دائم للكهرباء غير منقطع للطاقة و واقبات شدة التيار شيئان أساسيان إذا كان البلد الذي تعمل فيه يعاني من انقطاعات في التيار الكهربائي. ويجب استخدام أقلام الحبر الأخضر كلما قام فرد من فريق معالجة البيانات بتعديل على البيانات الخاصة بالاستبيان. تميز أقلام الحبر الأخضر هذه التغييرات عن البيانات الأصلية التي قامت الباحثة بتسجيلها (بالحبر الأزرق) وأية تغييرات يقوم بها فريق العمل الميداني (بالحبر الأحمر).

### حزم البرمجيات

تم تطوير البرامج النموذجية لمعالجة مسوحات الدورة الثالثة للمسوحات العنقودية متعددة المؤشرات في حزمة البرامج CSPro 2.6 و حزمة التحليل الاحصائي SPSS أما حزمة البرامج CSPro 2.6 الذي يستخدم لمعالجة كل من المسوحات والتعدادات السكانية فقد تم تطويره بالتعاون بين مكتب الولايات المتحدة للتعداد السكاني ومكتب المستشار الإقليمي ROC SerPro Ltda و Macro International. ويمكن تحميلها إلى جهاز الحاسوب مجاناً من الموقع الإلكتروني مكتب الولايات المتحدة للتعداد السكاني<sup>1</sup>. أما حزمة برامج التحليل الاحصائي SPSS، فهو حزمة برامج تجارية متوفر من خلال اليونيسف ومن خلال مزودي البرمجيات.

### مساحة المكتب

تلزم غرف منفصلة لإدخال البيانات ولتدقيق البيانات. أما غرفة إدخال البيانات فيجب أن تكون واسعة بصورة كافية حتى يكون لدى كل مدخل بيانات مساحة كافية لجهاز الحاسوب الخاص به أو بها وللاستبيانات في المكان الذي يشغله ويعمل فيه. ويجب ان توجد مكاتب وطاولات للعمل ومنافذ كهربائية كافية. يجب أن تكون الغرفة معتدلة الحرارة وإضاءتها جيدة وخالية من الغبار والرطوبة قدر المستطاع، ففي البلدان المعروفة بالمناخ الحار، يستلزم ذلك أن تكون الغرفة مكيفة ويجب توصيل مزود للكهرباء غير منقطع الطاقة إلى كل جهاز حاسوب، فإذا كانت فترات انقطاع الطاقة الكهربائية متكررة أو تدوم لفترة طويلة من الوقت، يجب وجود مولد كهرباء إذا لزم الأمر.

<sup>1</sup> عنوان الموقع الإلكتروني هو: <http://www.census.gov/ipc/www/CSPro/download/CSPro26.exe>

أما غرفة تدقيق البيانات فهي تخص منظم الاستبيان و المدققين الثانويين، ويجب أن تكون معتدلة الحرارة كذلك وجيدة الإضاءة، ويجب أن تكون هناك مساحة كافية لمدققي البيانات لمراجعة الاستبيانات. مثالياً، يجب أن تحتوي غرفة التدقيق على رفوف كافية أو خزائن لتخزين الاستبيانات بطريقة منظمة، فإذا لم يمكن تخزين الاستبيانات في غرفة التدقيق، يجب تخزينها في الجوار في متناول اليد حتى يسهل الوصول إليها كما تلزم الحاجة إليها في مختلف مراحل عملية المعالجة. كن حريصاً ألا تسيء تقدير مقدار الفراغ الذي قد يلزم لتخزين آلاف الاستبيانات التي ستكون لديك في المكتب في نهاية فترة العمل الميداني.

## تعديل البرامج النموذجية

كما هو موجز في الفصل الثالث فإن نموذج استبيان الدورة الثالثة للمسوحات العنقودية متعددة المؤشرات يجب أن يعدل حسب الحالة في كل بلد. مما يوجب تعديل البرامج النموذجية لإدخال البيانات والتدقيق والجدولة لتتماثل مع التغييرات التي أجريت على الاستبيان. وكلما زادت التغييرات التي تجري على نموذج الاستبيان، كلما ازداد الوقت الذي يلزم تخصيصه لتعديل وفحص البرامج. على سبيل المثال، إذا تمت إضافة أسئلة إلى الاستبيان، يجب إجراء إضافات متوافقة مع ذلك في برامج إدخال البيانات وفي تدقيقها وجدولتها.

سوف تكون هذه العملية أسهل بصورة جوهرية إذا ما تم الحفاظ على رقم السؤال في نموذج الاستبيان، فإذا تمت إضافة سؤال، يجب إضافة حرف إلى الترقيم الموجود (مثلاً، تم إدخال سؤال بين السؤالين WS4 و WS5، عندها يجب ترقيمه بالرقم WS4A).

و بالمثل، إذا تم حذف أسئلة، فيجب عدم إعادة ترقيم الأسئلة المتبقية. علاوة على ذلك، فإنه عند وضع رموز للفئات لإضافتها إلى نموذج الاستبيان، يجب إضافتها في نهاية القائمة الموجودة مع ترك الرموز الأخرى دون مسّها أو تغييرها. ويجب الانتهاء من تعديل برامج إدخال البيانات و برامج التدقيق قبل مرحلة الاختبار القبلي، فيمكن بذلك إدخال الاستبيانات في مرحلة الاختبار هذه وتدقيقها باستخدام البرامج.

وسوف يخدم هذا الإجراء هدفين اثنين، سوف يكشف المشاكل في نظام الترميز و الانتقالات الموجودة في الاستبيانات وسوف يكشف أيضاً أية أخطاء في البرامج. فحالما تنتهي عملية الاختبار القبلي والانتها من الاستبيانات، يجب إجراء التغييرات النهائية على البرامج. تزودك الأقسام التالية لهذا الفصل بإرشادات أساسية لتعديل قاموس البيانات النموذجي و التطبيقات النموذجية. يوجد ملخص توضيحي مستقل لمحتويات حزمة البرامج CSPro .

حتى و إن لم يتم إضافة أسئلة إلى الاستبيانات النموذجية، فإن قاموس البيانات النموذجي والتطبيقات تحتوي على متغيرات معينة يجب تحديثها (مثلاً، المدى المقبول لتاريخ المقابلة، المدى المقبول لرقم العنقود، الخ). وهذه المتغيرات هي بالضرورة محددة بالبلد ويجب إتقانها بواسطة أنت. وهكذا، حتى لو استخدمت بلدك نموذج الاستبيان، سيكون عليك تكييف البرامج النموذجية.

## قواميس البيانات

في الدورة الثالثة للمسوحات العنقودية متعددة المؤشرات، يتم جمع مجموعات الأسئلة المتصلة ببعضها في أقسام (مثلاً، وفيات الأمهات واستعمال موانع الحمل والتحصين)، والتي بعد ذلك يتم تجميعها في استبيانات (مثل، استبيان الأسرة، واستبيان المرأة، و استبيان الأطفال دون الخامسة). في حزمة البرامج CSPro، يتم استخدام قواميس لوصف هيكل البيانات: مجموعة المتغيرات (الأسئلة) المترابطة تكون سجل (قسم)، وتولف مجموعة السجلات مستوى (استبياناً)، يتم تخزينها معاً في ملف قاموس (بامتداد يدعى: dcf). وبالإضافة إلى قاموس البيانات، يتم استخدام شاشات لإدخال البيانات ترتبط بقواميس البيانات. وهناك في العادة شاشة إدخال واحدة لكل سجل.

يتم حفظ شاشات الإدخال على شكل ملف (ذو امتداد يدعى: fmf)، ويمكن تعديل ملفات dcf و fmf مباشرة. أفضل طريقة لعمل ذلك هي فتح ملفات الشاشات في برنامج CSPro سيتمكن هذا من الوصول إلى قاموس البيانات والشاشات معاً ويضمن تزامن التعديلات في كليهما معاً. ومما ننصح به الاحتفاظ بنسخة احتياطية لملفي القاموس و الشاشات النموذجيين كمرجعية.

هناك ثلاثة أنواع من استبيانات الدورة الثالثة للمسوحات العنقودية متعددة المؤشرات : الأسرة والنساء والأطفال دون سن الخامسة. أما استبيانات النساء والأطفال، فكل منها يرتبط بوحدة تحليل منفردة: المرأة والطفل على التوالي. ويتضمن استبيان الأسرة وحدتي تحليل: المنزل وأفراد الأسرة، ويتم تخزين جميع أنواع الاستبيانات في ملف تحت اسم mics 3.dcf وملف entry.fmf.

### المتغيرات التعريفية والمستويات

في برنامج CSPro، يجب أن يكون لكل استبيان سلسلة من المتغيرات التي تعرفه و تميزه. على سبيل المثال، يتم تعريف الأسرة بواسطة رقم عنقودها ورقم الأسرة في هذا العنقود، وتعرف المتغيرات التي تعرف استبياناً ما بالمتغيرات التعريفية. يدرج الجدول أدناه أنواع الاستبيانات و متغيراتها التعريفية.

رقم السطر	رقم الأسرة	رقم العنقود	الاستبيان
	HH2	HH1	الأسرة
LN	HH2	HH1	النساء
LN	HH2	HH1	الأطفال

كما نرى من الجدول، فإن للنساء والأطفال نفس المتغيرات التعريفية. وحيث أن كل فرد من أفراد الأسرة مدرج في سطر منفصل في قائمة الأسرة، فلن يكون لأي امرأتين أو طفلين أو امرأة و طفل من نفس الأسرة نفس رقم السطر. وهكذا، يجمعها مع رقم العنقود ورقم الأسرة، يصبح رقم السطر معرفاً بشكل منفرد للمرأة أو الطفل.

في أي قاموس منشأ داخل CSPro، يعرف المستوى مجموعة من المتغيرات التعريفية. في قاموس الدورة الثالثة للمسوحات العنقودية متعددة المؤشرات، هناك مستويين هما: الأسر والأفراد (أي المرأة المؤهلة والطفل المؤهل). تكون الأسر هي المستوى الأول، بينما تكون المرأة أو الطفل المستوى الثاني. و هذا التسلسل و الترتيب في المستويات طبيعي حيث أنه في استبيان



الدورة الثالثة للمسوحات العنقودية متعددة المؤشرات تنتمي كل امرأة وينتمي كل طفل إلى أسرة ما، بينما قد تتكون أسرة معينة من عدة نساء و عدة أطفال.

يتم تخزين استبيان النساء واستبيان الأطفال على نفس المستوى لأن كلا منها استبياناً ينطبق على فرد من أفراد الأسرة. يحتوي تطبيق إدخال البيانات على منطق يقوم بترك شاشات الإدخال المرتبطة بمتغيرات خاصة باستبيان الأطفال عند إدخال بيانات استبيان السيدة و ترك شاشات الإدخال المرتبطة بمتغيرات خاصة باستبيان النساء عند إدخال استبيان يخص طفلاً. وهكذا، فبالرغم من تخزين استبيانات النساء واستبيانات الأطفال كحالات في المستوى الثاني، إلا أنه ليس لهما متغيرات مشتركة سوى المتغيرات التعريفية.

## الأقسام

لقد تم تصميم قاموس البيانات لكي يعكس تكون استبيانات الدورة الثالثة للمسوحات العنقودية متعددة المؤشرات من أقسام. يتم تخزين كل قسم في سجله الخاص في mics3.dcf (باستثناء: قسمي قائمة الأسرة ودعم الأطفال الميتمين وغير المحصنين حيث أن لكل منها سجلين بسبب بنيتها الغير عادية)، ولكل سجل يوجد شاشة إدخال (أو اثنين في حالة قسم (سجل) خصائص الأسرة) مرتبط بها في entry.fmf. وهكذا، إذا كانت بلدك لا تستعمل قسماً معيناً، يمكن حذف السجل الخاص به وشاشة إدخاله (وحذف أية أوامر تشير إليه من تطبيق إدخال البيانات).

الأقسام الموجودة باستبيان الأسرة (مع رمز/ رموز) القسم المدرجة بين قوسين فهي: معلومات الأسرة (HH)، قائمة الأسرة (HL و TO)، التعليم (ED)، المياه والصرف الصحي (WS)، خصائص الأسرة (HC)، الشبكات المعالجة بالمبيدات الحشرية (TN)، دعم الأطفال الميتمين وغير المحصنين (OV و OR)، عمل الأطفال (CL)، تأديب الأطفال (CD)، الإعاقة (DA)، وفيات الأمهات (MM) ومعالجة الملح باليود (SI).

أما الأقسام الموجودة باستبيان المرأة فهي: خانة معلومات المرأة (WM)، وفيات الأطفال (CM)، التلقيح ضد الكزاز (TT)، صحة الأم والمولود (MN)، والزواج (MA)، ومتانة المسكن (ST) استعمال وسائل تنظيم الأسرة (CP)، وختان الأنثى (FG)، والمواقف تجاه العنف المنزلي (DV)، والسلوك الجنسي (SB) ومرض نقص المناعة البشرية المكتسبة / الايدز (HA).

وأما القوالب المتوفرة لاستبيان الأطفال فهي : خانة معلومات الأطفال دون سن الخامسة (VF)، وتسجيل الولادات والتعليم المبكر (BR)، تطور الأطفال (CE)، فيتامين أ (VA)، الرضاعة الطبيعية (BF)، العناية بالمرضى (CA)، والملاريا (ML)، والتحصين (IM) وقياسات الوزن والطول (AN).

## تسمية المتغيرات

يتم تسمية المتغيرات بقسم ما في الاستبيان حسب القسم الذي توجد به و رقم السؤال الذي تمثله. على سبيل المثال، يتم تخزين السؤال التاسع من قائمة الأسرة في متغير يدعى HL9. ويتم تقسيم بعض الأسئلة إلى جزئين أو أكثر، يعطي كل جزء منها حرفاً تعريفاً منفصلاً، ويتم تخزين كل جزء في متغير منفصل، وتتضمن أسماء هذه المتغيرات المنفصلة الحرف الذي يميز جزء السؤال. على سبيل المثال، السؤال 11 في قسم صحة الأم والمولود يتكون من جزئين. يتم

تخزين الجزء الأول من السؤال كمتغير باسم MN11A، أما الجزء الثاني فيتم تخزينه بالمتغير MN11.

تتكون بعض الأسئلة من جزأين أو أكثر للإجابة، ويتم تخزين هذه الأسئلة بمتغير منفرد، أما أجزاء الإجابة فيتم تعريفها بمتغير فرعي. وعندما تتعلق هذه الأسئلة بتواريخ، يتم إلحاق الأحرف d (ليشير إلى اليوم) و M (ليشير إلى الشهر) و Y (ليشير إلى السنة) إلى اسم المتغير الأساسي لتكوين اسم لهذا المتغير الفرعي. ففي السؤال 6 في قسم بيانات المرأة على سبيل المثال، فإن المطلوب إدخال يوم وشهر وسنة الميلاد، فيتم تخزين إجابتها في المتغير الرئيسي WM8 المكون من ثلاثة متغيرات فرعية هي: WM8d، WM8m، WM8y.

بعض الأسئلة لها بنية يكون فيها الجزء الأول من الإجابة هو وحدة الإجابة أما الجزء الثاني فهو الإجابة نفسها. يتم تخزين هذه الأسئلة تحت متغير واحد يتكون من متغيرين فرعيين هما الوحدة والإجابة نفسها، ويكون اسم المتغير الفرعي الذي يخزن وحدة الإجابة هو اسم المتغير الرئيسي مضاف إليه الحرف u (الذي يشير إلى الوحدة) بينما يتم تخزين المتغير الفرعي الذي يحتوي على الإجابة نفسها باسم المتغير الرئيسي مضاف إليه الحرف n (الذي يشير إلى العدد). على سبيل المثال، يسجل السؤال 13 في قسم صحة الأم والمولود كم مضى من الوقت بعد الولادة حتى تم إعطاء الطفل حليباً من صدر أمه، فإن الإجابة على هذا السؤال قد تكون بالساعات أو بالأيام. يتم في هذه الحالة تخزين الإجابة في المتغير wn13 المتكون من المتغيرين الفرعيين mn13n و mn13u.

### الأسئلة ذات الإجابات المتعددة والمتغيرات ذات الحروف الأبجدية

هناك عدد من الأسئلة في الاستبيان تسمح بتعدد الإجابات و يتم تمييز هذه الأسئلة بأن لها رموز إجابة عبارة عن حروف أبجدية (أي حروف من A - Z). في قاموس البيانات، يتم تخزين الإجابات لأسئلة تحتمل أكثر من إجابة واحدة في متغير من نوع أبجدي يكون طوله مساو لعدد الإجابات المحتملة، وهذه هي المتغيرات الوحيدة ذات الحروف الأبجدية الموجودة في القاموس. يتم تخزين المتغير الأبجدي كمتغير منفصل يتكون من متغيرات فرعية عددها هو عدد الإجابات المحتملة للسؤال الذي يمثله المتغير الرئيسي، و يتم تسجيل كل من المتغيرات الفرعية باسم المتغير الرئيسي مضاف إليه رمز الإجابة الذي يمثله المتغير الفرعي.

على سبيل المثال، السؤال الثاني من قسم صحة الأم والمولود يطلب تسجيل جميع الأفراد الذين تتلقى منهم الأم الرعاية في فترة الحمل قبل آخر ولادة لها. أما رموز الإجابات المحتملة فهي A و B و C و F و G و H و X و Y. وبذلك فإن المتغير mn2 يكون طوله ثمانية حروف و يتكون من ثمانية متغيرات فرعية هي: mn2a، mn2b، mn2c، mn2f، mn2g، mn2h، mn2x، mn2y.

### دليل الترميز

تستخدم القواميس النموذجية رموزاً نموذجية لإجابات معينة. سوف أناقش أولاً مصطلحات الترميز للمتغيرات الرقمية. الإجابة "أخرى" يتم ترميزها دائماً بواسطة العدد 6 و في الخانات على يساره الرقم تسعة، أما الإجابة "متناقض" فيرمز لها بالعدد 7 و في الخانات على يساره الرقم تسعة، وأما الإجابة "لا اعرف" فيرمز لها دائماً بالعدد 8 و الرقم تسعة في الخانات على

يسارها، في حين يتم ترميز الأسئلة التي لا تحتوي على إجابة (أي لم تقم الباحثة بتسجيل أية إجابة لسؤال قابل للإجابة) دائماً بالعدد 9 وفي الخانات إلى يساره الرقم 9، والأسئلة غير القابلة للإجابة عليها فتبقى فارغة. يلخص الجدول أدناه اصطلاحات الرموز القياسية :

الإجابة	طول المتغير			
	حروف	منزلة واحدة	منزلتين	ثلاثة منازل
أخرى	X	6	96	996
لا/لا شيء/لا احد	Y	na	na	na
غير متسق	na	7	97	997
لا أعرف	Z	8	98	998
مفقود	?	9	99	999
غير قابل للتطبيق	فارغ	فارغ	فارغ	فارغ

لأن الرموز من 6 - 9 يتم الاحتفاظ بها للاستخدام الخاص، فإن أي سؤال يتطلب أكثر من 6 إجابات يجب أن يمثلته متغير مكون من عددين مع وضع صفر على يسار الرمز (مثلاً 01 و02 و03 و04 و05 و06 و07 و96 و97 و98 و99).

بالنسبة للمتغيرات المكونة من حروف أبجدية، فيرمز للجواب "أخرى" دائماً بالرمز X، والإجابة "لا أعرف" يرمز لها دائماً بالرمز Z، ويرمز للقيمة الغير موجودة دائماً بعلامة الاستفهام (?) وتبقى الإجابات ل "غير قابلة للتطبيق" فارغة.

### مدى المتغيرات

إن معظم الأسئلة في استبيانات الدورة الثالثة للمسوحات العنقودية متعددة المؤشرات معرف لها مدى لقيم الإجابات. فمدى القيم للمتغيرات في قاموس mics3.dcf معرف. يقوم برنامج CSPro خلال عملية إدخال البيانات بالتحقق من أن أية قيمة يتم إدخالها في متغير هي ضمن المدى المعرف للمتغير. ويسمح برنامج CSPro بعدد كبير من تراوحدات المدى للإجابات لكل متغير. لذا الأسئلة ذات الإجابات الغير مسلسلة (مثل 1-8، 96، 98، 99) يجب تعريفها باستخدام تراوحدات المدى متعددة مثل 1-6، 96، 98، 99 بدلاً من 1-99، بينما تعريف مدى لقيم متغيرات القاموس مفيد للتحقق من التراوحدات البسيطة لمدى القيم فإن التراوحدات الأكثر تعقيداً أو التراوحدات الشرطية (مثلاً الاتساق بين اليوم والشهر في متغير يتعلق بالتاريخ) يجب التحقق منها في برامج الإدخال و التدقيق.

### تطبيق إدخال البيانات

إن تطبيق إدخال البيانات هو برنامج طويل ومعقد. و حيث أنه لا يمكن الخوض في تفاصيل برنامج الإدخال في هذا الفصل، فإن هذا القسم يركز على بعض المواضيع الأساسية في برنامج الإدخال.

### الانتقالات

تسمح استبيانات الدورة الثالثة للمسوحات العنقودية متعددة المؤشرات من استخدام عدد وافر من الانتقالات. الانتقالات هي تعليمات في الاستبيان للباحثة للانتقال من سؤال إلى سؤال آخر تاركة الأسئلة بين السؤالين دون طرحها على المبحوث/المبحوثة بناءً على إجابة السؤال الأول. يجب أن تتماثل الانتقالات في برنامج الإدخال مع الانتقالات في الاستبيان. الانتقالات المعرفة في برنامج إدخال البيانات تحدد مسار إدخال البيانات. و يغير برنامج CSPro مسار إدخال البيانات كلما تم استخدام أمر *Skip to* - انتقل إلى- أو *Skip to next* -انتقل إلى التالي.

## رسائل الخطأ

إذا قام مدخل البيانات بإدخال قيمة متغير لا تتوافق مع قيم لمتغيرات أخرى قد ادخلها مسبقاً ، فمن المفيد إظهار رسالة تدل على وجود خطأ في البيانات المدخلة – رسالة خطأ، ويجب أن توضح رسالة الخطأ هذه طبيعة المشكلة وأن تزود بأية معلومات يمكن أن تساعد على حلّ هذا التناقض. في برنامج CSPro يعرض الإجراء `errmsg` رسالة خطأ مع نصّ يعرفه مطور برنامج الإدخال كلما تم استدعاء هذه الإجراء . يتم ترقيم رسائل الخطأ لبرنامج إدخال البيانات وتخزينها في ملف `entry.mfg`، وقد أدرجنا نصوص و أرقام رسائل الخطأ المعرفة داخل البرنامج النموذجي للإدخال و تم إدراج أيضاً الأخطاء التي تؤدي إلى ظهور هذه الرسائل أثناء عملية الإدخال الملحق السابع وكذلك إرشادات عن كيفية حلها.

يجب عليك أن تراجع استبيانك لكي تحدد فيما إذا كانت أجوبة أي من الأسئلة المضافة تتطلب تحققاً من توافقها مع أجوبة أسئلة أخرى في الاستبيان، فإذا كانت تحتاج إلى تحقق، عليك إضافة الأوامر اللازمة في برنامج إدخال البيانات أو في برنامج التدقيق أو في كليهما. عند إضافتك فحص التناسق والتوافق، تأكد من إضافة رسالة خطأ متوافقة مع الفحص و توضحه إلى ملف رسائل برنامج إدخال البيانات أو إلى ملف رسائل برنامج التدقيق. علاوة على ذلك، إذا قمت بإضافة رسائل خطأ، فتأكد من عدم استعمالك رقم رسالة خطأ موجودة مسبقاً.

تتبع بعض رسائل الخطأ بأمر *reenter* -إعادة إدخال- و هو ما يتطلب إدخال قيمة صحيحة لذات المتغير التي تم إدخالها بصورة خاطئة . هذا الأمر يفرض على مدخل البيانات تصحيح الخطأ قبل مواصلة الإدخال. ولأن من المطلوب من مدخل البيانات أن يقوم أحياناً بإجراء تصحيحات، فإن ذلك يحتم وجود مراقبة دقيقة. فعندما تضيف رسالة خطأ خاصة بك، راعي بحرص فيما إذا كنت تريد أن تفرض على مدخل البيانات أن يقوم بتصحيح البيان الخطأ قبل أن يستمر في الإدخال. فإذا كنت تريد ذلك، قم بإتباع رسالة الخطأ بأمر *reenter* .

## المتغيرات التي تحتوي على حروف

يقوم تطبيق إدخال البيانات بالتحقق من أن المتغيرات ذات الحروف قد تم إدخالها بصورة صحيحة، فهو يقوم بعمل أربع عمليات تحقق على كل متغير يحتوي حروف. أولاً، يتحقق من أن القيمة التي تم إدخالها هي إحدى رموز الإجابة المدرجة في الاستبيان ( أي أن يتم التحقق من وجود القيمة في المدى المعرف). ثانياً يتحقق من أن الإجابات قد تم إدخالها حسب ترتيب الحروف الأبجدية ( أي ACG وليس GAC ) . ثالثاً يتحقق من أنه إذا تم إدخال رموز للإجابات " لا أعرف " أو "لا أحد " ( التي يرمز لها عادة بالرمز Y ) في الإجابة عندها لا يجب وجود أي إجابة أخرى ( أي، لا يسمح بالإجابة ACY ). رابعاً، يتحقق من أنه إذا تم إدخال رمز مفقود (؟) في الإجابة ، عندها لا توجد أية إجابة أخرى على السؤال ( أي ، لن يسمح بالإجابة ؟. (AC

يقوم تطبيق إدخال البيانات دائماً بإعادة ترتيب القيم المدخلة في متغيرات حرفية حتى يتم تخزين كل إجابة في الموقع الذي يعرفه متغيرها الفرعي. وللمتغير mn2 على سبيل المثال، فإن الإجابة ACG يتم إعادة ترتيبها إلى G C A، حيث يترك فراغ واحد بين الأحرف A و C و G وثلاثة فراغات بعد الحرف G .

### الإجراءات المعرفة بواسطة المستخدم

من الخصائص الحسنة في برنامج CSPro هو أنه يسمح للبرامج المصممة بواسطةه بأن تعرف إجراءاتها الخاصة. وهذه الإجراءات تعرف بواسطة المستخدم (مصمم البرنامج) ويمكن أن تكون مفيدة فهي على وجه الخصوص تسمح لمصمم البرنامج أن يتجنب إعادة كتابة الأوامر المستخدمة بصورة متكررة في برامج الإدخال أو التدقيق. وتكتب دائماً الإجراءات المعرفة من قبل المستخدم في بداية التطبيق أو البرنامج المصمم. ويحتوي تطبيق إدخال البيانات entry.app على 14 إجراء معرفة بواسطة المستخدم، ولا حاجة إلى تعديل هذه الإجراءات، إلا أن عليك فهم ما يمكنها أدائه إذا كان عليك أن تفهم تطبيق إدخال البيانات.

ويقوم الإجراء valid بالتحقق ما إذا كانت قيمة متغير ما هي إحدى القيم الخاصة: عدم اتساق، لا أعرف، إجابة مفقودة، أو غير قابلة للتطبيق. فإذا كانت قيمة المتغير غير قابلة للتطبيق، تقوم الإجراء natozero بتحويله إلى 0، مما يسمح بإضافته أو جمعه إلى متغير آخر (لإعطاء مثال على استخدامها، أنظر الإجراء cm9).

تستخدم الإجراءات السبعة التالية المعرفة من قبل المستخدم (zscr، zseval، dabs، zscoef، zgemth، zsanth) في استبيان الطفل وذلك لحساب قياسات الوزن و الطول التي توجد في نهاية قسم قياسات الوزن و الطول. يتم استخدام الإجراء agemthf لحساب عمر الطفل بالأشهر، ثم يتم استدعاء الإجراء zsanth. يقوم هذا الإجراء باستدعاء الإجراءات zseval و zscr و zspct، أما الإجراء zseval فيستدعي الإجراء zscoef يقوم الإجراء zspct باستدعاء الإجراء dabs. وسوف تجد هذه الإجراءات في متغيرات قياسات الوزن و الطول فقط، وحينها ستعرف بأنها تقوم بإحصاء والتحقق من قياسات الوزن و الطول.

أما الأوامر الموجودة في الإجراء agemth تقوم بحساب عمر الطفل بالأشهر، ولأن قياسات الوزن و الطول حساسة جداً للعمر، فيجب أن يحسب عمر الطفل اعتماداً على عمره بالأيام. تقوم الأوامر الموجودة بالإجراء أولاً بحساب عدد الأيام بين بداية العام ويوم ميلاد الطفل، ثم تقوم بعد ذلك بحساب عدد الأيام التي بين بداية العام وبين تاريخ إجراء المقابلة، وأخيراً تضاف عدد الأيام بين سنة الميلاد والسنة التي أجريت فيها المقابلة إلى عدد الأيام منذ بداية السنة حتى اليوم الذي أجريت فيه المقابلة. يكون عمر الطفل بالأيام هو الفرق بين هذين العددين. يتم بعد ذلك تحويل هذا الرقم إلى عمر الطفل بالأشهر وذلك بقسمته على 30.4375 (معدل عدد الأيام في الشهر على مدار أربع سنوات). و من أجل الدقة، يتم حساب عمر الطفل بالأشهر لأقرب عشرين.

تقوم الإجراءات vvalid و vdoi و vdob بالتحقق من تواريخ التطعيمات التي يتم إدخالها في قسم التحصين بحيث تكون صحيحة، لا تأتي بعد تاريخ المقابلة ولا تكون قبل تاريخ ولادة الطفل . أما الإجراء endmess (وهي اختصار الكلمة "نهاية الرسالة" "endmessage") فيعرض

رسالة في نهاية الاستبيان تسأل مدخل البيانات إذا كان يريد استعراض الاستبيان الحالي الذي قد تم إدخاله أم يريد المتابعة إلى الاستبيان الذي يليه. أخيراً، يقوم الإجراء alphack بعمليات التحقق من المتغيرات الحرفية كما هو مفصّل في القسم السابق.

### التواريخ و الترميز بالشهور منذ بداية القرن

تقوم البرامج النموجية (متضمنة تطبيق إدخال البيانات) ترميز معظم التواريخ لتكون بعدد الأشهر منذ بداية القرن (CMC). ترميز التواريخ بالأشهر منذ بداية القرن هو حساب عدد الأشهر منذ نهاية شهر ديسمبر منذ عام 1899 و حتى التاريخ المراد ترميزه. على سبيل المثال، يكون ترميز شهر يناير عام 1900 هو: 1 و ترميز شهر مارس عام 2000 هو 1203. ويتم حساب هذا الترميز لتاريخ معين كما يلي: اطرح 1900 من سنة التاريخ المعني، ثم اضرب الناتج من عملية الطرح ( عدد السنوات) في 12، ثم قم بإضافة رقم شهر التاريخ إلى ناتج عملية الضرب. على سبيل المثال يتم حساب الترميز لشهر مارس من عام 2000 بالمعادلة (2000-1900) x 12 + 3 .

يستخدم تطبيق إدخال البيانات أربع إجراءات لتبسيط عملية الترميز هذه. اثنين منها وهما setlb و setub يقومان بحساب الحدّين الأدنى والأعلى على التوالي لقيمة الترميز لشهر لتاريخ حدث معين. أما الإجراءين الأخيرين وهما adjlba و adjuba، فيقومان بتعديل الحدّين الأدنى والأعلى على التوالي لقيمة الترميز لتاريخ حدث معين مثل ( تاريخ ميلاد الطفل) بناءاً على عمر معطى. يلخص الجدول أدناه هذه الإجراءات. هذه الإجراءات تستخدم عموماً في حال عدم التأكد من شهر تاريخ الحدث.

<p>شهر، سنة، و حد أدنى لقيمة الترميز هم مدخلات هذا الإجراء. إذا كانت القيم المعطاة للمدخلين "month" و "year" و اللذين يمثلان الشهر و السنة قيم تصح لتمثيل شهر و سنة فإن الإجراء يقوم بحساب الترميز (كما وضح من قبل) و يكون ناتج و مخرج الإجراء هو قيمة الترميز. إذا كانت قيمة المدخل "year" و الذي يمثل السنة لا تصح لتمثيل سنة فعليا فإن ناتج و مخرج الإجراء هو القيمة المعطاة للمدخل "minimum" و الذي يمثل الحد الأدنى لقيمة الترميز و التي يحددها المستخدم. أما إذا كانت قيمة المدخل "month" و الذي يمثل الشهر لا تصح لتمثيل شهر فعليا فإن الإجراء يفترض أن الشهر هو يناير و يقوم بحساب قيمة الترميز على هذا الأساس و يكون هذا هو ناتج و مخرج الإجراء.</p>	<p>lcmc = setlb (month, year, minimum);</p>
<p>شهر، سنة، و حد أقصى لقيمة الترميز هم مدخلات هذا الإجراء. إذا كانت القيم المعطاة للمدخلين "month" و "year" و اللذين يمثلان الشهر و السنة قيم تصح لتمثيل شهر و سنة فإن الإجراء يقوم بحساب الترميز (كما وضح من قبل) و يكون ناتج و مخرج الإجراء هو قيمة الترميز. إذا كانت قيمة المدخل "year" و الذي يمثل السنة لا تصح لتمثيل سنة فعليا فإن ناتج و مخرج الإجراء هو القيمة المعطاة للمدخل "maximum" و الذي يمثل الحد الأدنى لقيمة الترميز و التي يحددها المستخدم. أما إذا كانت قيمة المدخل "month" و الذي يمثل الشهر لا تصح لتمثيل شهر فعليا فإن الإجراء يفترض أن الشهر هو ديسمبر و يقوم بحساب قيمة الترميز على هذا الأساس و يكون هذا هو ناتج و مخرج الإجراء.</p>	<p>ucmc = setub (month, year, maximum);</p>
<p>الحد الأدنى لقيمة ترميز تاريخ الحدث و هي ناتج استخدام الإجراء lcmc و الحد الأعلى لقيمة ترميز تاريخ الحدث و هي ناتج استخدام الإجراء ucmc و قيمة ترميز تاريخ المقابلة و العمر هم مدخلات هذا الإجراء و الذي يقوم بضبط و تعديل الحد الأدنى لقيمة الترميز (يكبره) بناء على العمر و تاريخ المقابلة لتقليل المدى بينه و بين الحد الأعلى. إذا كانت نتيجة حسابات الإجراء هي قيمة أكبر من أو تساوي الحد الأدنى المعطى و الذي يمثل المدخل lcmc و في نفس الوقت أقل من أو يساوي الحد الأعلى المعطى و الذي يمثل المدخل ucmc فيكون هذا هو ناتج و مخرج الإجراء. و إذا كانت نتيجة الحسابات هي قيمة أقل من الحد الأدنى المعطى lcmc فيكون ناتج الإجراء هو الحد الأدنى المعطى lcmc. أما إذا كانت نتيجة الحسابات هي قيمة أكبر من الحد الأعلى المعطى و الذي يمثل المدخل ucmc فيكون ناتج و مخرج الإجراء هو -1.</p>	<pre>t = adjlba (lcmc, ucmc, di, di, age); if t &lt; 0 then   e = errmsg (2171); else   lcmc = t; endif;</pre>
<p>الحد الأدنى لقيمة ترميز تاريخ الحدث و هي ناتج استخدام الإجراء lcmc و الحد الأعلى لقيمة ترميز تاريخ الحدث و هي ناتج استخدام الإجراء ucmc و قيمة ترميز تاريخ المقابلة و العمر هم مدخلات هذا الإجراء و الذي يقوم بضبط و تعديل الحد الأعلى لقيمة الترميز (يقبله) بناء على العمر و تاريخ المقابلة لتقليل المدى بينه و بين الحد الأدنى. إذا كانت نتيجة حسابات</p>	<pre>t = adjuba (lcmc, ucmc, di, di, age); if t &lt; 0 then   e = errmsg (2171); else   ucmc = t; endif;</pre>

الإجراء هي قيمة أكبر من أو تساوي الحد الأدنى المعطى و الذي يمثل المدخل lcmc و في نفس الوقت أقل من أو يساوي الحد الأعلى المعطى و الذي يمثل المدخل ucmc فيكون هذا هو ناتج و مخرج الإجراء. و إذا كانت نتيجة الحسابات هي قيمة أكبر من الحد الأعلى المعطى ucmc فيكون ناتج الإجراء هو الحد الأعلى المعطى ucmc. أما إذا كانت نتيجة الحسابات هي قيمة أقل من الحد الأدنى المعطى و الذي يمثل المدخل lcmc فيكون ناتج و مخرج الإجراء هو -1.

## استلام الاستبيانات من الميدان

عندما تصل الاستبيانات لعنقود ما من الميدان، يجب على منظم الاستبيانات التحقق من عدد الاستبيانات على صفحة ضبط العنقود، فإذا كانت هناك استبيانات مفقودة، على منظم الاستبيانات الاتصال بفريق العمل الميداني لمعرفة إمكانية العثور عليها. فإذا لم يكن ذلك ممكناً، يجب على فريق العمل الميداني إعادة تسجيل الاستبيانات المفقودة إذا كان ذلك ممكناً (أي، إذا كان فريق العمل الميداني ما يزال قريباً من العنقود). وإذا كان لم يكن ممكناً وكان الاستبيان المفقود هو استبيان أسرة، يجب على منظم الاستبيانات أن يضيف استبياناً فارغاً (باستثناء البيانات التعريفية) مع الرمز للنتيجة بالرقم 6 لاستبيانات العنقود. وإذا كان الاستبيان المفقود هو استبيان امرأة أو استبيان طفل، فعلى منظم الاستبيانات تغيير إجمالي عدد الاستبيانات المكتملة على صفحة الغلاف في استبيان الأسرة الذي تنتمي إليها المرأة أو ينتمي إليها الطفل.

يجب تسجيل وصول الاستبيانات من الميدان بواسطة منظم الاستبيانات على نموذج تتبّع العنقود. يتوفر نموذج تتبّع العنقود بالنمطيين الإلكتروني والورقي. يمكن إيجاد مثال عن النمط الورقي للنموذج في نهاية هذا الفصل، بينما يتوفر الإصدار الإلكتروني للنموذج في قائمة المشرف. قبل استقبال أية استبيانات من الميدان، يجب إدراج العناقيد بترتيب تصاعدي على نموذج تتبّع العنقود. وحين استقبال استبيانات العناقيد، يمكن عندها تسجيل معلومات عن كل استبيان في الخانات المخصصة له. بالإضافة إلى ذلك، يجب أن يتأكد منظم الاستبيانات مما يلي لكل عنقود:

- أن جميع الاستبيانات المدرجة في صفحة ضبط العنقود موجودة.
- أن الاستبيانات الخاصة بالأسرة والنساء والأطفال لتلك الأسرة موضوعة مع بعضها.
- أن الاستبيانات مرتبة حسب رقم الأسرة ترتيباً تصاعدياً.
- جميع استبيانات النساء موضوعة مع بعضها بالترتيب التصاعدي لرقم سطر كل امرأة، متنوعة بجميع استبيانات الأطفال كذلك بالترتيب التصاعدي لرقم سطر كل طفل.
- أن رموز جميع المعلومات الجغرافية ومعلومات المقابلات مكتملة في خانة معلومات الأسرة التي تكون في بداية كل استبيان للأسرة.



حالما تكتمل عملية التحقق وتحلّ أية مشاكل قد تطرأ، يجب على منظم الاستبيانات تسجيل عدد استبيانات الأسر والنساء والأطفال على نموذج تتبّع العناقيد. وعليه أيضاً أن يقوم بإنشاء صفحة ملخص للعنقود. تكون صفحة ملخص العنقود هي من الورق السميك و تحوي الاستبيانات وعليها رقم العنقود بالخط العريض. ويجب أن يكتب عليها أيضاً عدد الاستبيانات في العنقود كما هو مبين أدناه.

حال إنشاء ملخص للعنقود، على المشرف إعطاء استبيانات هذا العنقود إلى مدخل بيانات (لاحظ بأن مدخل البيانات الموكل إليه هذا العنقود يجب أن يقوم بإدخال جميع استبيانات العنقود). يجب على منظم الاستبيانات بعد ذلك أن يدوّن اسم مدخل البيانات على نموذج تتبّع العناقيد.

_____	إجمالي استبيانات الأسر
_____	إجمالي استبيانات الأسر المكتملة
_____	إجمالي استبيانات النساء
_____	إجمالي استبيانات النساء المكتملة
_____	إجمالي استبيانات الأطفال
_____	إجمالي استبيانات الأطفال المكتملة

كلما تم إخراج استبيانات من موقع تخزينها للقيام بإدخال بياناتها أو تدقيقها، يجب إعادتها دائماً بتخزينها بنفس الطريقة التي كانت عليها مسبقاً وإعادتها إلى نفس موقعها السابق، فسوف تجعل عملية التعامل مع الاستبيانات بطريقة نظامية ستساعد في سرعة إنجاز خطوات معالجة البيانات بطريقة فعالة.

بعد إدخال بيانات كل عنقود على نموذج تتبّع العناقيد، يجب أن يقوم المشرف بمراجعة الاستبيانات والتحقق من أنها جاهزة لإدخال بياناتها. ندرج في الجدول أدناه عمليات التحقق التي يجب القيام بها.

وحالما تنتهي عملية التحقق وتحلّ جميع المشاكل التي قد تظهر، يقوم المشرف بإعطاء العنقود إلى مدخل البيانات.

### إدخال البيانات الرئيسي

من الأفضل أن تبدأ عملية إدخال البيانات بعد أن يبدأ العمل في الميدان، فإن هذا سيسمح لك بتحديد وتصحيح الأخطاء التي قد تقمن بها الباحثات أو فرق العمل الميداني، ويمكن التعرف على مشاكل جدية قد يسهوا المشرف الميداني عن ملاحظاتها في الوقت المناسب لإعادتها إلى الكادر الميداني وتصحيح الأخطاء.

والمشرفون الميدانيون مسؤولون عن التحقق من الاستبيانات من حيث اكتمالها وتوافقها وعن تصنيف أية إجابات لم يستطع مجري المقابلة وضع رمز لها و عليه فإن التدقيق المكتبي يكون عملية بسيطة. حالما يصل نموذج البيانات من عنقود واحد إلى المركز الرئيسي، يجب البدء بعملية إدخال بياناته.

### التدقيق المكتبي

1. التأكد من أن جميع استبيانات الأسرة مرتبة ترتيباً تصاعدياً حسب رقم الأسرة في العنقود.
2. التأكد من أن جميع البيانات الجغرافية و بيانات المقابلة قد اكتملت في خانة بيانات الأسرة المعيشية في بداية كل استبيان للأسرة.
3. يجب التحقق من أهلية كل فرد من أفراد الأسرة من قائمة أفراد الأسرة. وحتى يكون الفرد مؤهلاً للاستبيان الخاص بالمرأة، يجب أن تكون أنثى بعمر يتراوح من 15 إلى 49 سنة، وحتى يكون الفرد مؤهلاً لقسم عمل الأطفال، يجب أن يكون الفرد بعمر يتراوح بين 5 سنوات إلى 14 سنة (أو ضمن المدى العمري الخاص بالدولة)، وحتى يكون الطفل مؤهلاً لأقسام صحة الطفل، يجب أن يكون أقل من خمس سنوات. لكل فرد مؤهل، يجب وضع دائرة حول رمز الأهلية لرقم السطر الخاص بهم أو رقم سطر أمهاتهم أو المتكفلين بهم (HL6 - HL8 في قائمة أفراد الأسرة). أما لباقي أفراد الأسرة، يجب إبقاء خانة الأهلية الخاصة بهم فارغة في قائمة أفراد الأسرة و سيتم إدخال القيمة 00 لهم. قم باتباع الإرشادات الموجودة في دليل التدقيق (الملحق السادس) فيما يتعلق بالرسائل من 0101 - 0132 لحل أية مشكلة تتعلق بالأهلية.
4. العدد الإجمالي للنساء المؤهلات وللأطفال المؤهلين المسجل في خانة معلومات الأسرة يجب أن يتطابق مع العدد المؤهل المدون في نموذج قائمة أفراد الأسرة.
5. إجمالي عدد النساء المؤهلات والأطفال المؤهلين المسجل في خانة معلومات الأسرة يجب أن يتطابق عدد استبيانات النساء والأطفال لتلك الأسرة.
6. عدد المقابلات المكتملة للنساء وللأطفال المسجل في خانة بيانات الأسرة يجب أن يتطابق مع عدد استبيانات النساء و استبيانات الأطفال المدون بها رمز النتيجة 1 (مكتمل في تلك الأسرة).
7. التأكد من أن جميع استبيانات النساء موضوعة مع بعضها بترتيب أرقام أسطر النساء ترتيباً تصاعدياً (WM4)، ومتبوعة باستبيانات جميع الأطفال بترتيب أرقام أسطرهم ترتيباً تصاعدياً.

يتم إدخال البيانات بواسطة برنامج مصمم على CSPro المسمى entry.ent يقوم هذا البرنامج بأداء مهمتين أساسيتين، أولاًهما، لكل أسرة يضمن أن جميع الاستبيانات الخاصة فقط بهذه الأسرة قد تم إدخالها البيانات وأن جميع بيانات هذه الاستبيانات فقط قد تم إدخالها. وبالتحديد يبدأ برنامج إدخال البيانات بإدخال جميع البيانات من استبيان الأسرة، وحالما يتم إدخال هذه البيانات، يقوم البرنامج بمراجعة قائمة أفراد الأسرة ويحدد من هم المؤهلون لاستبيانات النساء. بعد تحديد هؤلاء السيدات، يطلب برنامج إدخال البيانات لكل سيدة مؤهلة الاستبيان الخاص بها حسب الترتيب التصاعدي لرقم سطر السيدة في قائمة أفراد الأسرة. ثم يقوم تطبيق إدخال البيانات بتنفيذ نفس الإجراء لأفراد الأسرة المؤهلين لاستبيان الأطفال. عند إدخال استبيانات الأسرة جميعها، يتم تخزين بيانات الأسرة بأكملها وينتقل البرنامج إلى الأسرة التي تليها في العنقود إن وجدت.

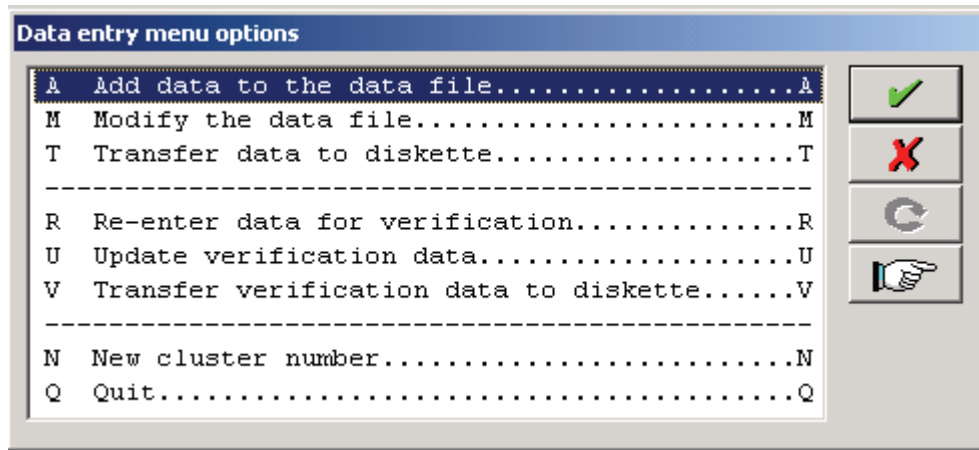
لأن بيانات جميع استبيانات الأسرة ونسائها المؤهلات وأطفالها المؤهلين تحفظ بعد إدخالها جميعها فقط، لا يجب لمدخل البيانات أن يترك حاسوبه في منتصف عملية إدخال بيانات الأسرة، فقبل أخذ فترة استراحة أو التوقف عن العمل في ذلك اليوم، يجب أن يتم إدخال جميع الاستبيانات الخاصة بأسرة ما. وفوق ذلك فإنه ينصح بنسخ البيانات على حاسوب المشرف أو على قرص لين كإجراء أمني. علاوة على ذلك، يجب على المشرف كل مساء أن يقوم بنسخ محتويات C:\mics وجميع أدلتها الفرعية على أداة تخزين ثانوية. سوف يسمح هذا الحفظ الآمن أن يستعيد المشرف البيانات فيما إذا تحطم جهاز حاسوبه.

بالإضافة إلى ضبط الاستبيانات التي يجب إدخالها، يقوم برنامج إدخال البيانات بتنفيذ الانتقالات الموجودة باستبيان ما بدقة. هذا يعني، بأن التطبيق سيطلب فقط الإجابات على الأسئلة التي يجب أن تطرح بناء على أجوبة الأسئلة التي تسبقها. على سبيل المثال، إذا تم إدخال القيمة 2 في المتغير Cm1 (أي لم يسبق للمرأة أن أنجبت)، سيقوم برنامج إدخال البيانات بعد ذلك بطلب قيمة للمتغير ma1 تاركاً جميع المتغيرات بين هذين المتغيرين و التي تخص فقط المرأة التي سبق وأنجبت.

أما المهمة الثانية لبرنامج إدخال البيانات فهو تقليل أخطاء إدخال البيانات. يقوم برنامج إدخال البيانات بهذه المهمة بواسطة إجراء عمليات تحقق أثناء إدخال البيانات، فإذا كانت القيمة التي يتم إدخالها لسؤال خارج مدى القيم لهذا السؤال في الاستبيان، أو إذا تم تقصي بعض التناقضات الأساسية الأخرى، يقوم البرنامج بإظهار رسالة خطأ وتتطلب من مدخل البيانات حلّ هذا التناقض قبل المتابعة. أما التناقضات الأكثر تعقيداً والتي سوف يبطن حلها عملية إدخال البيانات بصورة جوهريّة، لا يتم التحقق منها أثناء إدخال البيانات بل يتم التحقق منها بدلاً من ذلك أثناء التدقيق الثانوي للبيانات.

## قائمة إدخال البيانات

تبسّط قائمة إدخال البيانات مهمة إدخال البيانات. لقد تم إنشاء قائمة إدخال البيانات بواسطة برنامج مصمم على CSPro وهو entry\_menu.bch (لا حاجة بك إلى تعديل هذا البرنامج). يبين النص التالي قائمة إدخال البيانات.



تستخدم الخيارات A و M و T فقط إذا كان مدخل إدخال البيانات يقوم بإدخال ملف البيانات الرئيسي، وتستخدم الخيارات R و U و V فقط إذا كان مدخل البيانات يقوم بإدخال ملف بيانات التحقق، أما الخيارين N و Q فيمكن استخدامه في كلا الحالتين.

يسمح الخيار A لمدخل البيانات بإضافة حالات إلى ملف البيانات الرئيسي، ويسمح الخيار M له بتعديل الحالات الموجودة في ملف البيانات الرئيسي، أما الخيار T فيسمح لمدخل البيانات بنسخ ملف البيانات الرئيسي المكتمل في قرص لِيّن حتى يتم نقله إلى جهاز حاسوب المشرف. ويسمح الخيار R لمدخل البيانات بإضافة حالات إلى ملف بيانات التدقيق، ويسمح له الخيار U بتعديل الحالات الموجودة على ملف التحقق، أما الخيار V فيسمح له بنسخ ملف البيانات التحقق المكتمل إلى قرص لِيّن حتى يتم نقله إلى جهاز حاسوب المشرف.

يسمح الخيار N لمدخل البيانات بتغيير رقم العنقود حتى يقوم بإدخال بيانات عنقود آخر، أما الخيار Q فيقوم بالخروج من قائمة إدخال البيانات.

### بنية الدليل على حاسوب إدخال البيانات

في حاسوب إدخال البيانات، يتم إدراج جميع الملفات والبرامج في الدليل C:\mics\CSPro أو أحد الأدلة الفرعية له.

يتم تسمية الأدلة الفرعية باسم data و dicts و entry و veri.

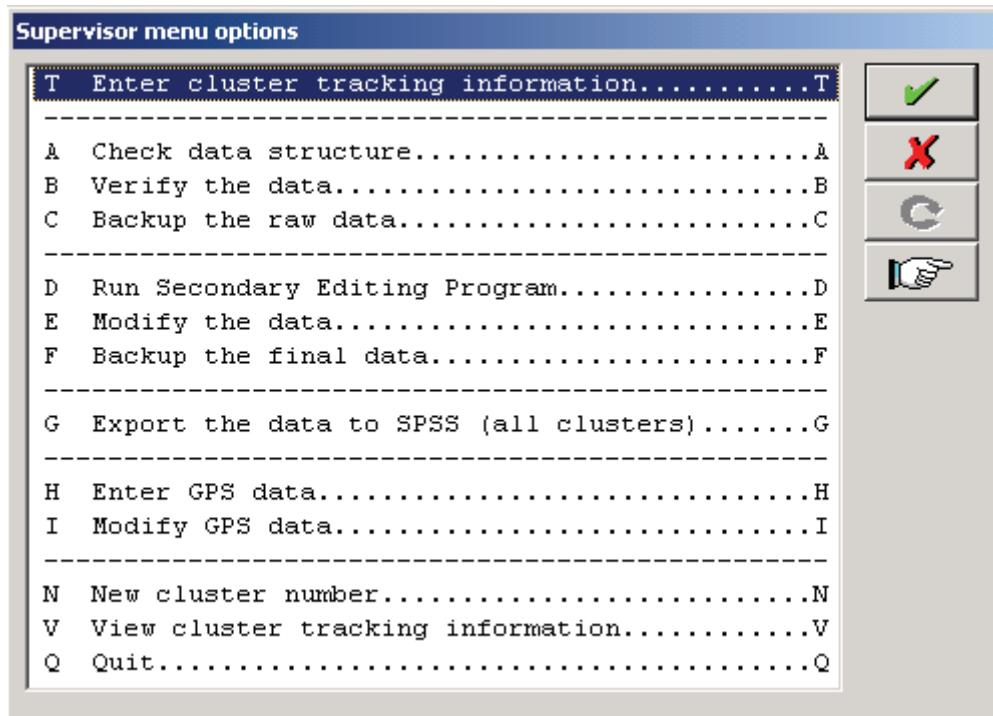
يحتوي دليل data على ملفات البيانات الرئيسية التي تم إدخالها إلى الحاسوب، ويحتوي الدليل dicts على جميع قواميس البيانات، ويحتوي الدليل entry على برنامج إدخال البيانات والبرنامج الذي يقوم بإنشاء قائمة إدخال البيانات، أما الدليل veri فيحتوي على أية ملفات بيانات تحقق يتم إدخالها على الحاسوب.

### بنية دليل حاسوب عامل إدخال البيانات:

		MICS
		CSPRO
	DATA	
	DICTS	
	ENTRY	
	VERI	
دليل البيانات الرئيسي		
دليل قواميس البيانات		
دليل برنامج إدخال البيانات		
دليل بيانات التحقق		

### قائمة المشرف

قائمة المشرف تساعد المشرف على إدارة نظام معالجة بيانات الدورة الثالثة من المسوحات العنقودية متعددة المؤشرات. وتفتح مباشرة بعد تنفيذ البرنامج المصمم على CSPro وهو super\_menu.pff الذي يطلب أولاً رقم العنقود المراد العمل عليه، وحالما يتم إدخال رقم العنقود تفتح القائمة، يبين النص أدناه قائمة المشرف يتبعها ملخص عن كل خيار.



الخيار T- إدخال معلومات تتبّع العنقود

يعرض هذا الخيار النموذج الإلكتروني لتتبع العنقود حتى يتمكن مشرف معالجة البيانات من إدخال المعلومات الخاصة بالعنقود. يجب أن يتم إدخال بيانات العنقود عندما يتم استقبال الاستبيانات لهذا العنقود على الفور، وعندما يتم إعطاء العنقود لمدخل البيانات الرئيسية، وعندما يتم إعطاء العنقود لمدخل البيانات ليقوم بإدخال بيانات التحقق. يتم تحديث المعلومات الأخرى في نموذج تتبّع العنقود الإلكتروني بصورة تلقائية عندما يواصل المشرف العمل عبر قائمة المشرف.

### الخيار A: التحقق من بنية البيانات

يتحقق هذا الخيار من بنية ملف بيانات العنقود بواسطة تنفيذ التطبيق `check.bch`. بعد التحقق من بنية ملف البيانات، يقوم `check.bch` بتكوين ملف مخرجات (يتم عرضه تلقائياً على الشاشة) تلخص أعداد كل نوع من أنواع الاستبيانات وتبين عدد الاستبيانات المكتملة. يجب على مشرف معالجة البيانات التحقق من المعلومات الموجودة بهذا الملف في نموذج تتبّع العنقود و أنهما متوافقين، فإذا وجد اختلاف بين الملف و النموذج، على مشرف معالجة البيانات إيجاد المشكلة (مثلاً، نسي مدخل البيانات أن يدخل أسرة ما) ثم يحلّ المشكلة عن طريق التحقق الدقيق من استبيانات هذا العنقود. حال حلّ المشكلة (سواء بتحديث ملف البيانات أو بتحديث نموذج تتبّع العنقود)، يجب على مشرف معالجة البيانات في تنفيذ البرنامج `check.bch` مرة أخرى، وعندما ينتج أمر `check.bch` نفس أرقام الاستبيانات المدونة في نموذج تتبّع العنقود فقط يستطيع مشرف البيانات عندها إعطاء الاستبيانات إلى مدخل بيانات ثاني لإدخال بيانات التحقق.

### الخيار B: التحقق من البيانات

يقوم هذا الخيار بمقارنة ملف البيانات الرئيسي للعنقود بملف بيانات التحقق باستخدام أداة المقارنة الموجودة في البرنامج `CSPPro`. إذا وجدت أية فروق بين ملفات البيانات فسوف تظهر هذه الفروقات على الشاشة. يجب طباعة هذه النتيجة وإعطاءها إلى مدخلي البيانات المسؤولين عن إدخال بيانات هذا العنقود. وبالعامل سويًا، على مدخلي البيانات الرجوع إلى الاستبيانات وتحديد القيمة الصحيحة لكل تناقض موجود بين الملفين، بعد تحديد القيم الصحيحة للتناقضات يجب على كل مدخل أن يحدث الملف الخاص به تبعاً لهذه التصحيحات. في هذه المرحلة يجب مقارنة الملفات مرة أخرى، وعندما لا تبقى هناك أية فروقات بين الملفين، يمكن بعدها مواصلة عملية معالجة العنقود.

### الخيار C: حفظ نسخة احتياطية من ملف البيانات الخام

يقوم هذا الخيار بعمل نسخة احتياطية من ملف البيانات الخام وذلك بنسخ ملف البيانات الرئيسية المحقق في دليل فرعي على حاسوب المشرف. يجب تنفيذ هذا الخيار بعد اكتمال عملية التحقق من البنية و التحقق من البيانات المدخلة وقبل إجراء التدقيق الثانوي. يتم حفظ نسخة احتياطية من البيانات الخام كمرجعية للبيانات قبل إجراء التدقيق الثانوي عليها.

### الخيار D: تنفيذ برنامج التدقيق الثانوي

يقوم هذا الخيار بالتحقق من التناقضات المعقدة وذلك بتنفيذ التطبيق `editing.bch`. يتم عرض الملف المخرج من هذا البرنامج على الشاشة ويجب طباعته إذا كان يتضمن رسائل خطأ، فإذا

كانت تحتوي على مثل رسائل الخطأ هذه، يجب إعطاء قائمة بالتناقضات واستبيانات العنقود إلى مدقق ثانوي.

يقوم المدقق الثانوي بالاعتماد على معرفته بالاستبيان وعلى كتيب التدقيق في الملحق السابع، بكلّ أي من التناقضات المدرجة. عندما ينتهي المدققون الثانويون من أعمالهم، يقومون بإعادة قائمة التناقضات والإجراءات التي ينبغي عملها إلى مشرف معالجة البيانات الذي يقوم بتطبيق التغييرات المطلوبة (انظر الخيار E أدناه)، ثم يعيد مشرف معالجة البيانات تنفيذ editing.bch، فإذا لم تظهر رسائل خطأ، يمكن عندها متابعة معالجة العنقود، أما إذا ظهرت رسائل الخطأ، فيجب إعطاء قائمة بها مع الاستبيانات إلى مدقق ثانوي لعمل تدقيق آخر. في بعض الحالات، يمكن أن يعتبر المدقق الثانوي رسائل الخطأ مقبولة وأنه لا حاجة لإجراء أي تصحيح آخر على البيانات. فإذا كانت الرسائل التي تظهر هي الرسائل التي اعتبرها المدقق الثانوي مقبولة فقط، عندها يعتبر عمل المدقق الثانوي قد اكتمل.

### الخيار E: تعديل البيانات

يقوم هذا الخيار بتنفيذ تطبيق entry.ent حتى يتمكن مشرف معالجة البيانات من إجراء التغييرات اللازمة برأي المدققين الثانويين. بعد استخدام هذا الخيار، يجب أن يعود مشرف معالجة البيانات إلى الخيار D لكي يتأكد من أن التغييرات التي تمت على البيانات قد قامت بتصحيح التناقضات كما هو مرغوب و أيضا من عدم وجود أية تناقضات جديدة.

### الخيار F. حفظ نسخة احتياطية من البيانات النهائية

يقوم هذا الخيار بنسخ ملفات البيانات النهائية في الدليل الفرعي final على حاسوب مشرف معالجة البيانات. يتم في مرحلة لاحقة ضم البيانات المخزنة في هذا الدليل في ملف واحد ثم تحويلها إلى حزمة التحليل الاحصائي SPSS.

### الخيار G: تحويل البيانات إلى حزمة التحليل الاحصائي SPSS.

يقوم هذا الخيار بضم كل ملفات البيانات في الدليل الفرعي final في ملف واحد ثم يقوم بتحويل هذه البيانات إلى حزمة التحليل الاحصائي عن طريق تنفيذ التطبيق export.bch. ينتج هذا التطبيق أربعة ملفات من نمط ASCII تحوي البيانات نفسها وملف وصفي لتركيبية و قاموس البيانات لكل ملف نصي من نوع SPSS. إحدى الملفات النصية تحتوي على بيانات الأسر، ويحتوي الآخر على أفراد الأسرة، ويحتوي ثالث على النساء ويتضمن الأخير الأطفال في الأسرة.

### الخيار H: إدخال البيانات الجغرافية.

يسمح هذا الخيار لمشرف معالجة البيانات بإدخال البيانات الجغرافية وذلك بتنفيذ تطبيق gpsentry.ent. وبعكس برنامج إدخال البيانات الرئيسية و الذي يتطلب تحديد عنقود معين في البداية لإدخال بياناته، فإن هذا التطبيق يسمح لمشرف معالجة البيانات بإدخال البيانات الجغرافية لأي عدد يريده من العناقيد. يطلب هذا التطبيق من مشرف معالجة البيانات أن يقوم بإدخال البيانات الجغرافية مرتين للتأكد من صحة الأرقام المدخلة.

## الخيار I: تعديل البيانات الجغرافية.

يسمح هذا الخيار لمشرف معالجة البيانات بتعديل البيانات الجغرافية وذلك بتنفيذ التطبيق .gpsentry.

## الخيار V: عرض معلومات تتبّع العنقود.

يقوم هذا الخيار بعرض المعلومات المخزنة في نموذج تتبّع العنقود لكل العناقيد الموجودة.

## الخيار N: اختيار عنقود جديد

يقوم هذا الخيار بتغيير رقم العنقود حتى يتمكن مشرف معالجة البيانات من البدء بمعالجة عنقود جديد.

## بنية الأدلة على حاسوب مشرف معالجة البيانات

تستخدم بيانات الجولة الثالثة من المسوحات العنقودية متعددة البيانات بنية دليل محددة على أجهزة حواسيب إدخال البيانات وعلى حاسوب المشرف على معالجة البيانات، وسيتم توضيح بنية حواسيب إدخال البيانات في الرسم التالي. على حاسوب المشرف، يتم تخزين جميع الملفات والبرامج التي تتعلق بحزمة التحليل الإحصائي SPSS في الدليل C:\mics\spss أو أحد أدلته الفرعية و يتم تخزين جميع الملفات والبرامج المتصلة ببرنامج CSPro في الدليل C:\mics\spss أو أحد أدلته الفرعية. يتم تسمية الأدلة الفرعية للدليل CSPro بالأسماء .weights ، super ، raw ، final ، gps ، export ، entry ، dicts ، backup.

## بنية الأدلة حاسوب المشرف:

	MICS
	CSPRO
دليل حفظ النسخ الاحتياطية من الملفات	BACKUP
دليل قواميس البيانات	DICTS
دليل تطبيق إدخال البيانات	ENTRY
دليل تطبيق تحويل البيانات	EXPORT
دليل البيانات المدققة النهائية	FINAL
دليل تطبيق إدخال البيانات الجغرافية GPS	GPS
دليل البيانات الخام	RAW
دليل تطبيق تدقيق المشرف	SUPER
دليل تطبيق أوزان العينات	WEIGHTS

يحتوي دليل backup على نسخ احتياطية لملفات البيانات التي تم التحقق من بنية بياناتها و المتحقق بياناتها المدخلة لكن غير المدققة بعد. ويحتوي دليل dicts على جميع قواميس البيانات. ويحتوي دليل entry على تطبيق إدخال البيانات والتطبيق الذي يقوم بإنشاء قائمة إدخال البيانات. ويحتوي دليل export على البرامج التي تستخدم لتحويل البيانات من برنامج CSPro إلى حزمة التحليل الإحصائي SPSS. يحتوي دليل final على نسخ احتياطية لملفات البيانات التي تم التحقق من بنيتها و من بياناتها وتدقيقها. ويحتوي دليل raw على ملفات البيانات التي تم نقلها من أجهزة إدخال البيانات. أما الدليل super، فيحتوي على التطبيقات التي تقوم بوظيفة التحقق من



البنية وصحة البيانات المدخلة وإجراء التدقيق الثانوي لها والتطبيق الذي يقوم بإنشاء قائمة المشرف. وأما دليل weights، فيحتوي على جدول يقوم بحساب أوزان العينات.

## التحقق من البنية

من الضروري أن تكون البيانات سليمة من الناحية الهيكلية. يفرض برنامج إدخال البيانات بأوامره أن تكون هيكلية البيانات صحيحة، إلا أنه لا يقوم بالتحقق من كل شيء فهذا يبطئ عملية الإدخال، لذلك فإنه من الضروري تنفيذ برنامج التحقق من البنية بعد الانتهاء من إدخال البيانات الرئيسية. يقوم برنامج التحقق من البنية بالتأكد من أن عدد الاستبيانات في ملف البيانات يطابق عدد الاستبيانات التي وصلت من الميدان، ويقوم بإجراء بعض الفحوصات الإضافية على بنية الاستبيانات الفردية.

يدعى برنامج التحقق من البنية check.bnh. سوف نركز في هذا القسم على ماذا يفعل البرنامج أكثر من كيف يعمل البرنامج. إن هذا البرنامج معقد و شرحه ليس هينا. وأفضل طريقة لفهم منطقية البرنامج هي بدراسة بحرص بعد أن تستوعب ماذا يفعل البرنامج. نبيّن أدناه عينة من مخرج لبرنامج check.bch.

MICS Data Structure Check  
Cluster: 3

Households	Women	Children
Total Comp Incomp	Eligible	Interviewed
HH12 FOUND	HH13 FOUND	HH14 FOUND
2 1 1	5 5 4 4	4 4 3 3

الخانة الأولى هي ملخص عن إجمالي عدد الأسر ورموز لنتيجة المقابلة (Households)، أما الخانة الثانية فهي المخرجات التي تمثل نتائج ل 4 أرقام تخص النساء المؤهلات (Women). هذه الأرقام هي قيم hh12 و hh13 و اللذان يشيران إلى عدد النساء المؤهلات وعدد النساء اللواتي تمت مقابلتهن طبقاً لخانة بيانات الأسرة. وأما الأرقام المدرجة تحت عمود found "موجود"، فتشير إلى عدد استبيانات النساء وعدد استبيانات النساء التي اكتملت في ملف البيانات على التوالي.

وأما الخانة الثالثة من المخرجات فتمثل نتائج تخص الأطفال المؤهلين دون سن الخامسة (Children)، فالأرقام الموجودة تحت العمودين hh14 و hh15 هي أعداد الأطفال المؤهلين دون سن الخامسة والأطفال المؤهلين الذين تمت مقابلتهم طبقاً لخانة بيانات الأسرة. وأما الأرقام المدرجة تحت عمود found "موجود"، فتشير إلى عدد استبيانات الأطفال دون سن الخامسة وعدد استبيانات الأطفال التي اكتملت في ملف البيانات على التوالي.

يجب أن تتم طباعة مخرجات البرنامج check.bch بواسطة المشرف ومقارنة البيانات التي تحتويها مع نموذج تتبّع العنقود، فإذا كان هناك فرق بين أعداد الاستبيانات، يجب على المشرف وعامل إدخال البيانات عندئذ استخدام قائمة الأخطاء واستبيانات العنقود لتحديد ما الذي سبب المشكلة في البنية. بعد التعرف على هذه الأسباب، يجب أن يقوم مدخل البيانات بتصحيحها. يجب بعد ذلك إعادة تنفيذ برنامج التحقق من البنية للتحقق من أنه قد تم حل المشكلة دون أن تخلق

مشكلة جديدة. عندما تتطابق الأرقام التي يخرجها برنامج check.bch مع تلك الأرقام الموجودة في نموذج تتبّع العنقود يمكن وقتها البدء بإدخال بيانات التحقق.

يقدم التطبيق check.bch أيضاً قائمة بكل الأسر في العنقود، ويظهر رقم الأسرة ورمز نتيجة المقابلة مع عدد استبيانات النساء المؤهلات وعدد الأطفال المؤهلين إذا اكتملت المقابلة مع الأسرة (أي، أن يكون رمز نتيجة الأسرة مساوياً للعدد 1). نبيّن أدناه عينة من المخرجات لإحدى الأسر.

MICS Data Structure Check

Household: 1

Result: 1

Women		Children	
Eligible	Interviewed	Eligible	Interviewed
HH12 FOUND	4	HH13 FOUND	3
	4	HH14 FOUND	2
	3	HH15 FOUND	1
	3		2
	2		1
	1		1

يمكن أن يكون إدراج الأسر ضمن جدول أمر مفيد للتعرف على مصدر أية مشكلة على مستوى العنقود. لنفترض مثلاً بأن صفحة تتبّع العنقود قد أدرجت 20 أسرة في العنقود، ولكن تم العثور على 19 أسرة منها، تستطيع عندها أن تحدّد أي الأسر هي التي لم يتم إدخال بياناتها.

## التحقق

يتم عمل التحقق من البيانات التي تم إدخالها مرتين بواسطة تطبيق المقارنة في برنامج CSpro. يدعى تطبيق المقارنة compare.cmp ، ويحتوي على قائمة بكل متغيرات قاموس البيانات المستخدم. و من طريقة عمل البرنامج أن فقط المتغيرات التي يتم اختيارها من قاموس البيانات هي التي تدخل في إجراء المقارنة. وهناك متغير واحد فقط لا ينبغي اختياره ضمن عملية المقارنة وهو المتغير HH16 والذي يمثل كود مدخل البيانات (فمن المفروض أن مدخل البيانات الرئيسية ليس هو مدخل بيانات التحقق) ولا ينبغي استبعاد أية متغيرات أخرى لأن وجود اختلافات في المتغيرات المستبعدة دون العلم بها يضعف من دقة البيانات.

يقوم تطبيق المقارنة بمقارنة ملف إدخال البيانات الرئيسي (الذي تم نسخه إلى حاسوب المشرف) مع ملف إدخال البيانات المدققة (التي تم نسخها إلى قرص لين أو مشغل شبكة إذا كانت الأجهزة متصلة بشبكة واحدة) و يعرض قائمة بالاختلافات إن وجدت. فإذا لم توجد أية فروقات، على المشرف أن يقوم بعمل نسخة احتياطية من ملف البيانات الخام ثم ينتقل إلى عملية التدقيق الثانوي.

أما إذا وجدت أية فروقات، فيجب طباعة قائمة بهذه الفروقات وإعطائها إلى عاملي إدخال البيانات. يقوم عاملي إدخال البيانات عندئذ باستخدام قائمة الفروقات واستبيانات العنقود للتحقق من كل فرق ثم التسجيل على القائمة اسم الملف الذي يحتاج إلى تصحيح. بعد التحري عن جميع الفروقات، يقوم عاملي إدخال البيانات بتصحيح الأخطاء في ملفاتهم، ثم يقومان بعد ذلك بإعادة نسخ الملفات على أقراصهم اللينة (أو على الشبكة)، ثم تتم مقارنة ملفاتهم مرة أخرى، وتستمر هذه العملية حتى تصبح الملفات متماثلة.

## التدقيق الثانوي

أثبتت التجربة أن التناقضات البسيطة يمكن التعرف عليها وتصحيحها أثناء عملية إدخال البيانات. أما الأخطاء و التناقضات الأكثر تعقيداً، فيجب أن يتم حلها بتحقيق دقيق من البيانات في الاستبيانات. معالجة هذه التناقضات من الأفضل أن تتم كخطوة منفصلة عن عملية الإدخال حتى لا تبطئ من عملية الإدخال، مع طبع تقرير بالأخطاء لتحديد أي منها يجب تصحيحه، وتعرف هذه الخطوة بالتدقيق الثانوي.

يدعى برنامج التدقيق الثانوي editing.bch ويقوم بتنفيذ قائمة طويلة من فحوصات عدم التوافق ( مثلاً توافق الأعمار وتواريخ الميلاد ) ويقدم قائمة بالتناقضات الموجودة في ملف البيانات الخام . يجب على مشرف معالجة البيانات أن يقوم بطباعة هذه القائمة وأن يعطيها مع استبيانات العنقود إلى أحد المدققين الثانويين. يقوم المدقق الثانوي بعد ذلك بمراجعة قائمة الأخطاء والإجابات الموجودة على الاستبيانات. باستخدام إرشادات التدقيق ( الموجودة في الملحق السابق ) واستناداً إلى معرفة المدقق الثانوي بالاستبيانات، فإنه يقوم بتدوين التصحيح على قائمة الخطأ أو يكتب عليها أنه لا يوجد أي إجراء يمكن عمله إزاء ذلك. بعدما يقوم المدقق الثانوي بمراجعة كل رسائل الخطأ، يعيد قائمة الأخطاء التي تم وضع الملاحظات عليها إلى مشرف معالجة البيانات. بعد ذلك، يقوم مشرف معالجة البيانات بتنفيذ التغييرات المقترحة في ملف البيانات الخام، ثم يعيد تنفيذ برنامج التدقيق. و عندما لا يصدر برنامج التدقيق أية رسالة خطأ، يمكن الاستمرار في عملية مراجعة العنقود.

هناك ثلاثة خطوات مهمة جداً في عملية التدقيق. الأولى أنه لكل خطأ، يجب على المدقق الثانوي فحص الاستبيان المعني بدقة شديدة. الثانية، أنه يجب على المدقق الثانوي دائماً أن يراجع إرشادات التدقيق قبل أن يضع حلولاً للمشكلة. والخطوة الثالثة هو أن عملية التدقيق يجب أن تتكرر حتى لا تبقى هناك أية رسالة خطأ.

عندما تنتهي الخطوة الثالثة، يمكن لمشرف معالجة البيانات حفظ نسخة احتياطية من ملف البيانات المدققة، وتعتبر هذه البيانات الآن نظيفة ويمكن استخدامها لإنشاء ملفات تحليل.

## تحويل البيانات إلى حزمة التحليل الإحصائي SPSS

عند الانتهاء من المرحلة الأولى لمعالجة للبيانات ، سيكون هناك ملف بيانات نظيف لكل عنقود في العينة. وبينما تتم هذه المرحلة من معالجة البيانات في برنامج CSPro، فإن المرحلة الثانية من معالجة البيانات تتم بصورة رئيسية باستخدام حزمة التحليل الإحصائي SPSS. لذلك فإن الخطوة الأولى في المرحلة الثانية من معالجة البيانات هي تحويل البيانات من برنامج CSPro إلى حزمة التحليل الإحصائي SPSS. يتم ذلك باستخدام خيار *Export the data to SPSS* "تحويل البيانات إلى حزمة التحليل الإحصائي SPSS" من قائمة المشرف.

عندما اختيار هذا الخيار، يتم ضم جميع ملفات البيانات الموجودة في الدليل الفرعي final (أي، جميع ملفات البيانات التي تم التحقق منها وفحصها وتدقيقها) في ملف بيانات واحد يسمى all.dat . يتم بعد ذلك تحويل ملف البيانات هذا إلى حزمة التحليل الإحصائي SPSS بواسطة تطبيق export.bch . يقوم هذا التطبيق بإنشاء أربعة ملفات بيانات من نوع ASCII هي ( mych.dat ، myhh.dat ، mygl.dat ، mywm.dat ) وأربعة ملفات وصفية لهيكل هذه البيانات لبرنامج SPSS هي ( mych.sps ، myhh.sps ، mygl.sps ، mywm.sps ) في

الدليل C:\mics\spss . وبينما تقوم الأوامر الموجودة في ملفات وصف البيانات هذه بقراءة ملفات البيانات المحولة من برنامج CSPRO من نوع ASCII في حزمة التحليل الإحصائي SPSS، إلا أنها لن تحفظها في صورة نهائية من نوع SPSS. ولجعل ملفات وصف البيانات تقوم بتخزين البيانات على نسق SPSS، يجب إضافة الأمر التالي إلى نهاية كل ملف وصفي.

Save outfile = filename.sav'.

ويجب استبدال كلمة filename بالأحرف hh و hl و wm و ch اعتماداً على محتويات ملف البيانات. عند تعديل الأمر بصورة مناسبة وإضافته إلى كل ملف من ملفات وصف البيانات، عند تنفيذ ملفات وصف البيانات يتم إنشاء ملفات بيانات SPSS هي hh.sav ، hl.sav ، wm.sav ، ch.sav.

### إنشاء ملف تحليل

يساعد ترميز المتغيرات أثناء المرحلة الأولى من معالجة البيانات في تسهيل عملية إدخال البيانات، و لكن هذا الترميز ليس مفضل لتحليل البيانات التي تم جمعها، لذلك فإن أول خطوة بعد نقل البيانات إلى حزمة التحليل الإحصائي SPSS هي إعادة ترميز المتغيرات لجعل التحليل أسهل وأكثر فاعلية. تعرف هذه الخطوة بإنشاء ملف تحليل. يقوم هذا القسم بتفصيل الخطوات التي تتعلق بإنشاء ملفات للتحليل للدورة الثالثة من المسوحات العنقودية متعددة المؤشرات. و يمكن تنفيذ تطبيقات الجدولة على هذه الملفات كما يمكن توزيعها على الباحثين .

### إعادة ترميز المتغيرات

تقوم التطبيقات المصممة على الحزمة الإحصائية SPSS وهي makehl.sps و makech.sps و makewm.sps بإعادة ترميز بعض المتغيرات الموجودة. تستخدم هذه التطبيقات في إعادة ترميز المتغيرات المستخدمة بصفة متكررة في الجداول و حفظها في متغيرات جديدة بالترميز الجديد أما المتغيرات الغير مستخدمة تكرر في الجداول فيتم إعادة ترميزها أثناء تنفيذ تطبيقات الجداول.

يتم إعادة ترميز المتغيرات باستخدام أوامر الحزمة الإحصائية SPSS ولن نقوم بمناقشتها هنا. على أية حال، فإن هناك طريقة تستخدم بصورة متكررة يجب أن نوضحها وهي إعادة ترميز المتغيرات و إعطاء المتغير القيمة 0 أو 100. يستخدم هذا الترميز لأغراض التمثيل الصحيح فقط. عند عرض جدول ما في حزمة التحليل الإحصائي SPSS فإنه يتم عرض كل القيم التي يأخذها المتغير الذي تم اختياره. هناك العديد من الجداول في خطة الجدولة التي تهتم فقط بقيمة واحدة للمتغير، فعند إعادة ترميز هذا المتغير و استبدال هذه القيمة بالرقم 100 و استبدال أية قيمة أخرى بالعدد 0 فعند حساب المتوسط لهذا المتغير الجديد تكون النتيجة هي نسبة الإجابة بالقيمة 100 لهذا المتغير الجديد و هي القيمة المراد تمثيلها في البداية للمتغير الذي تم إعادة ترميزه.

على سبيل المثال، المتغير ha1 يقوم بتسجيل ما إذا كانت المرأة قد سمعت بمرض الإيدز. فإذا كانت المرأة قد سمعت بمرض الإيدز يأخذ المتغير القيمة 1، ويأخذ المتغير القيمة 2 إذا لم تكن قد سمعت بهذا المرض. نحن مهتمون بعرض النسبة المئوية للنساء اللواتي سمعن بمرض الإيدز. و

لذلك في التطبيق makewm.sps، يتم إعادة ترميز المتغير ha1 و حفظه في المتغير الجديد aids. يأخذ المتغير aids القيمة 100 في حال كانت المرأة قد سمعت بمرض الإيدز وبعكس ذلك تأخذ القيمة 0. يكون المتوسط للمتغير aids هو النسبة المئوية للنساء اللواتي سمعن بمرض الإيدز. ولاستيعاب ذلك، تابع المثال التالي.

10	النساء اللواتي سمعن بمرض الإيدز
20	إجمالي عدد النساء
$50 = 100 \times 20 \div 10$	النسبة المئوية للنساء اللواتي سمعن بمرض الإيدز
$50 = 100 \times 20 \div 10 = 20 \div 100 \times 10 = 20 \div (0 \times 10 + 100 \times 10)$	متوسط المتغير aids

## حساب وإضافة أوزان العينات

إذا تم تصميم العينة من البداية بألا تكون ممثلة تمثيلاً وطنياً أي بناءً على الطبيعة المختلفة للمناطق ( أو المقاطعات ) ، فإن العينة الوطنية لم يتم اختيارها على أساس نظرية تناسب الاحتمالية مع الحجم (PPS) Probability Proportional to Size أيضاً إذا تم التمثيل بناءً على بعض العوامل الأخرى ( مثلًا، حضر/ ريف أو عشوائيات/ غير عشوائيات) وأخذت في الاعتبار بعض العوامل الأخرى فإن العينة لا تكون ذاتية التوزيع و يجب إضافة الأوزان عند حساب التقديرات و النسب، يجب التأكد من ان كل عينة فرعية و التي تمثل منطقة معينة تشارك فقط في العينة الكلية كما لو كانت العينة ممثلة تمثيلاً وطنياً.

إذا لم تكن العينة ذاتية التوزيع، فيجب حساب الأوزان و إضافتها إلى ملفات التحليل. يتم ذلك باستخدام ملف الجداول weights.xls و التطبيقات weights\_table.sps و weight.sps و weights\_merge.sps المصممة على SPSS. يستخدم دليل الجداول لحساب الأوزان، ويتضمن جدولين هما calculation -حسابات- و output -ناتج. أما الجدول calculations فيقوم بإجراء العمليات الحسابية لحساب الأوزان، وأما الجدول output، فيحتوي على أوزان العينة بأرقام العناقيد و وجود الجدول بهذا الهيكل يسهل من قراءته في SPSS. و يقوم تطبيق weights\_table.sps بتقديم البيانات اللازمة لحساب أوزان العينة، ويضيف التطبيق weights\_merge.sps أوزان العينة إلى ملفات التحليل، أما البرنامج weights.sps الذي لن تقوم أنت بتنفيذه مباشرة مطلقاً، فهو يصف هيكل البيانات في الجدول output.

و يمكن تقسيم عملية حساب و إضافة الأوزان إلى سبعة خطوات و هي:

الخطوة الأولى – قم بتعديل عدد الصفوف في جدول calculations و جدول output بحيث يكون هناك صف واحد لكل عنقود موجود في المسح. بعد أن تضيف أو تحذف صفوف بناءً على عدد العناقيد الموجودة في العينة، تأكد من أن هذا التعديل لا يغير صف totals –إجمالي- في الجدول calculations.

الخطوة الثانية – قم بإدخال الأوزان التي تم تحديدها وقت تصميم العينة إلى ملف weights.xls. فإذا كانت أوزانك تتباين عبر العناقيد ضمن منطقة (أو مقاطعة) معينة، عليك أن تدخل بيانات كل من العمود cluster sampling fraction - معامل العينة للعنقود- و العمود stratum sampling fraction - معامل العينة للطبقة (أو المقاطعة) و التي زدك بها خبير تصميم عينة المسح. وإذا كانت الأوزان تختلف حسب المنطقة (أو المقاطعة) ولكن ليس حسب العناقيد في هذه الطبقة (أي إذا كانت العينة ذاتية التوزيع على مستوى المنطقة أو المقاطعة)، قم بإدخال القيمة 1

في العמוד cluster sampling fraction و اكمل بيانات العמוד stratum sampling fraction وذلك حسب القيم التي يحددها خبير تصميم عينة المسح.

الخطوة الثالثة – قم بتعديل تعريف الطبقة (أو المقاطعة) في الأسطر من 3 - 10 من برنامج weights\_table.sps. تفترض البرامج النموذجية أن المنطقة يحددها جميع خصائص هذه المنطقة و الإقليم التابعة له (أي حضر/ ريف) ، وأن هناك أربعة أقاليم (يجب تعديل البرنامج لكي يعكس تقسيمة المناطق في عينة المسح الخاصة بالدولة).

الخطوة الرابعة – قم بتنفيذ البرنامج weights\_table.sps.

الخطوة الخامسة – قم بنسخ البيانات في الجدول ثم إلصاقها في الجدول calculations في الملف weights.xls. وعندما تنتهي من هذه الخطوة، سوف يقوم ملف weights.xls بحساب أوزان العينة.

الخطوة السادسة – قم بحفظ الجدول output كملف بامتداد (\*.csv) باسم weights.csv في الدليل c:\mics.weights.

الخطوة السابعة – قم بتنفيذ برنامج weights\_merge.sps، وعندما تنتهي من الخطوة السابعة، تحقق من عدم وجود رسائل خطأ بعد تنفيذ البرنامج و افتح ملف التحليل وتأكد من أن الأوزان قد تم دمجها في الملف بصورة صحيحة.

### حساب وإضافة معامل الثروة

يوجد في خطة جدولة المسوحات العنقودية متعددة المؤشرات معامل الثروة للأسرة كمتغير. يتم حساب معامل الثروة هذا بواسطة البرنامج wealth.sps و الذي يقوم بإنشاء ملف بيانات wealth.sav يحتوي على متغيرات البيانات التعريفية، و متغير يحتوي على مؤشر ثروة الأسرة، و متغير يمثل معامل الثروة للأسرة. يقوم برنامج wealth.sps أولاً بإنتاج تكرارات لجميع متغيرات الأسرة التي تتعلق بالثروة أو بالأصول، ثم يقوم بعد ذلك بإعادة ترميز المتغيرات الخاصة بممتلكات و أصول الأسرة و ممتلكات و أصول الفرد إلى متغيرات جديدة مختلفة. يقوم البرنامج بعد ذلك باستخدام طريقة تحليل العوامل (تحليل المكونات الرئيسية بالتحديد) لحساب مؤشر ثروة لكل أسرة. وأخيراً، يستخدم البرنامج مؤشر الثروة هذا لإنشاء معاملات لثروة الأسرة (معامل الثروة) ثم يحفظها في ملف بيانات من نوع SPSS. إن اختيار المتغيرات التي يجب تضمينها في تحليل العوامل أمر يحتاج إلى دقة، ولا يجب القيام به دون الرجوع إلى نتيجة التكرارات التي يقوم بحسابها البرنامج wealth.sps، وسيتم الخروج بقرارات عن أي المتغيرات التي يجب تضمينها و عن كيفية إعادة تسجيل هذه المتغيرات في ورشة العمل الثالثة. و بعد حساب معامل الثروة، سيقوم برنامج wealth\_merge.sps بإضافته إلى ملفات التحليل الخاصة بك. تحقق من عدم وجود رسائل خطأ بعد تنفيذ البرنامج و افتح ملف التحليل وتأكد من أن معامل الثروة قد تم دمجها في الملف بصورة صحيحة.

### إضافة البيانات الجغرافية GPS

ستسجل بعض المسوحات العنقودية متعددة المؤشرات قراءة للموقع الجغرافي لعناقيدها خلال العمل الميداني، و التي يمكن استخدامها لتحديد و إضافة بيانات جغرافية أخرى (على سبيل المثال: سقوط الأمطار). فإذا كان مسح البلد يسجل قراءات الموقع الجغرافي، هذا يتطلب دمجها

في ملفات التحليل. يتم استخدام برنامجين مصممين بواسطة الحزمة الإحصائية SPSS هما برنامج gps.sps وبرنامج gps\_merge.sps. أما برنامج gps.sps الذي لن تقوم أبداً بتنفيذه مباشرة، فيصف بنية ملف البيانات gps.dat (الذي يقوم تطبيق إدخال البيانات gpsentry.bch المصمم بواسطة البرنامج CSPro بإنشائه). وإذا كان هناك تعديلات في قاموس البيانات الجغرافية gps.dic، فيجب تعديل البرنامج gps.sps لكي يتوافق مع التعديلات التي أجريت على قاموس البيانات. أما برنامج gps\_merge.sps فيدمج القراءات الجغرافية في ملف التحليل ولن يكون عليك تعديل هذا البرنامج. لذا فدمج القراءات الجغرافية في ملفات التحليل، قم بتنفيذ برنامج gps\_merge.sps ثم تحقق من عدم وجود رسائل خطأ بعد تنفيذ البرنامج و افتح ملف التحليل وتأكد من أنه قد تم دمج القراءات الجغرافية في الملف بصورة صحيحة.

## الجدولة

يشرح الفصل الثامن بالتفصيل عملية تحليل البيانات وتحضير التقارير. تنفيذ برامج إنشاء الجداول هو خطوة رئيسية في هذه العملية. هذا الكتيب مصحوب بنموذج لمنظومة الجداول (الفصل الثامن والملحق السادس) وبرامج إنشاء الجداول المصممة بواسطة الحزمة الإحصائية SPSS. هناك برنامج لكل جدول المسوحات العنقودية متعددة المؤشرات، واسم كل برنامج هو الحرف t متبوعاً برقم الجدول في منظومة الجداول. على سبيل المثال، يقوم البرنامج t1.sps بإنشاء الجدول رقم 1 في منظومة الجداول.

قبل تنفيذ برامج الجدولة إعداد تقرير ما، من الجوهرى إنتاج مجموعة من توزيعات تكرارات (غير محتسبة الأوزان) لكل متغير في ملف البيانات. يجب التحقق من التكرارات للتأكد من عدم وجود قيم غير عادية، كتلك التي تكون خارج مدى الإجابات، وتلك الإجابات التي تبدو بعيدة التصديق للأسئلة. على سبيل المثال، فإن الإجابة "53" عن السؤال الذي يتعلق بعدد الساعات التي عمل بها الطفل في أعمال المنزل خلال الأسبوع الماضي تبدو عالية بصورة بعيدة عن التصديق ومحددة جداً، ويجب تدوين البيانات التعريفية لمثل هذه الحالات والتحقق من ملف البيانات مقارنة مع الاستبيانات الأصلية.

## مراجعة البرامج النموذجية

يجب مراجعة كل برنامج من برامج إنشاء الجداول بعناية، للتأكد أن كل المتغيرات التي تم استخدامها في هذه البرامج موجودة بالفعل في ملف البيانات. فإذا كان هناك متغير غير موجود، قم بالتحقق فيما إذا كان المتغير ذا أهمية رئيسية أم ثانوية. إذا كان المتغير ذو أهمية رئيسية و غير موجود في ملف البيانات، فيجب استبعاد الجدول بأكمله أو الطلب من المحلل أن يعيد تصميم الجدول. أما إذا كان المتغير الغير موجود ذو أهمية ثانوية، قم بحذفه اسم المتغير من برنامج إنشاء الجداول، ثم عمل التعديلات الأخرى الضرورية تبعاً لذلك.

يجب التحقق بعناية أيضاً من إجراءات إعادة ترميز المتغيرات، فإذا وجد في الاستبيانات الخاصة بالدولة أية متغيرات تحتوي على فئات ترميز غير نموذجية، يجب تعديل إجراءات إعادة الترميز لمثل هذه المتغيرات، كذلك إذا كان الاستبيان يحتوي على متغيرات مضافة من قبل الدولة المنفذة فيجب إعادة ترميزها إذا كان سيتم استخدامها في أي جدول.

و يجب أيضاً التحقق من إجراءات دمج البيانات إذا كان الاستبيان الخاص بك يستخدم متغيرات تعريفية أخرى غير الموجودة في الاستبيانات النموذجية حيث يوجد عدد من إجراءات الدمج في

برامج الجدولة و التي لن تنفذ إلا إذا تم استخدام كل المتغيرات التعريفية لتمييز الحالات في ملفات البيانات فعلاً.

## استخدام أوزان العينة

استخدام الأوزان في برامج الجدولة هو إجراء مباشر إلا إذا تم استخدام أمر التجميع الأمر aggregate. فإذا كان الهدف من استخدام الأمر aggregate هو تجميع الحالات لحساب البسط والمقام، فيجب استخدام الأوزان قبل تنفيذ الأمر aggregate و لا يجب استخدام الأوزان مرة أخرى في الملف الناتج عن استخدام أمر التجميع aggregate لأنه قد تم استخدام الوزن بالفعل قبل تنفيذ هذا الأمر.

على سبيل المثال، الجدول HH.1 يعرض معدل الاستجابة، و هو ما يتطلب قسمة متغيرين في نفس الجدول. أحد الطرق لحساب ذلك هو إنشاء ملف تجميعي يحتوي على عدد الأسر في العينة، المنازل المأهولة والأسر التي تمت مقابلتها سيحتوي ملف التجميع هذا على حالة واحدة لكل فئة من قيم متغير التجميع break variable الذي يتم تحديده (مثل: ريف / حضر).

يجب استخدام الأوزان عند إنشاء ملف التجميع لحساب بسط موزون (عدد الأسر التي تمت مقابلتها) ومقام موزون (عدد المنازل المأهولة)، وعندما يتم إنشاء ملف التجميع، يحسب معدل الاستجابة لكل فئة من قيم متغير التجميع بقسمة البسط على المقام.

أما إذا كان الهدف من استخدام الأمر aggregate هو حساب احصائيات على الحالات الفردية في ملف البيانات، فيتم استخدام الأوزان بعد تنفيذ أمر aggregate. على سبيل المثال، يحتوي الجدول 4 على معلومات عن نسبة الأسر التي يوجد بها على الأقل طفل واحد دون سن الخامسة عشر.

هذه المعلومات غير موجودة في ملف بيانات الأسرة، إلا أنه يمكن حسابها وذلك بتجميع بيانات ملف قائمة أفراد الأسرة حسب متغيرات التجميع المحددة و هي رقم العنقود ورقم الأسرة. يتم استخدام الأوزان بعد عملية التجميع لأن المراد هنا هو حساب النسبة المرجحة للأسر التي يوجد بها على الأقل طفل واحد أقل من 15 سنة و ليس حساب العدد المرجح لعدد الأطفال الأقل من 15 سنة في كل أسرة.

## أمر التضمين include

يقوم البرنامج tables.sps بتنفيذ برامج الجدولة كلها. يحتوي البرنامج على الأمر include لكل برنامج من برامج الجدولة لتنفيذه. فإذا صادف البرنامج tables.sps خطأ في برنامج ما من برامج الجدولة المراد تنفيذها (التي يتم تنفيذ أمر include عليها)، فإن البرنامج tables.sps سيتوقف.

لذلك يجب عدم استخدام البرنامج tables.sps إلا بعد أن يتم التحقق من جميع برامج الجدولة واختبارها. تأكد أيضاً من حذف أي أمر include يمكن أن ينفذ برنامج جدول لا تستخدمه.

يفرض أمر include أربعة قيود على البرامج التي يقوم الأمر بتنفيذها. القيد الأول هو أن كل أمر يجب أن يبدأ في العمود الأول من البرنامج. هذا القيد يحد من عدد الفراغات في بداية أسطر



الكتابة، لكن السطر يمكن أن يحتوي على فراغ إذا بدأ بالعلامة "+" . يوضّح الأمر التالي استخدام العلامة "+" .

```
do if (cage >= 6 and cage <= 9).
+ compute solids = 0.
+ if (BF3G = 1) solids = 100.
end if.
variable labels solids "Solid foods".
```

القيد الثاني الذي يفرضه أمر Include هو أن الأمر الذي يكتب على عدة أسطر، يجب أن يكون العمود الأول من الأسطر التالية لأول سطر فارغاً. يوضح المثال التالي أمراً للأسطر متعددة تسير حسب هذا القيد.

```
add files
/file=*
/file='tmp6.sav'.
```

لاحظ أن الأوامر الفرعية في السطرين الثاني والثالث تحتوي على فراغين في البداية (بينما يفرض هذا القيد إبقاء العمود الأول فارغاً، إلا أنه تم إبقاء أول عمودين فارغين لتتوافق مع طريقة كتابة بقية برامج المسوحات العنقودية متعددة المؤشرات).

أما القيد الثالث والرابع اللذان يفرضهما أمر include، فيتمثلان بأن وضع علامة نهاية الأمر شيء اختياري وأن العلامة (\*) في العمود الأول من السطر تشير إلى أن السطر يحتوي على ملاحظات أو توضيحات و ليس أوامر للتنفيذ، ولا يؤثر أي من هذين القيدين على البرامج الخاصة بإنشاء الجداول.

بالإضافة إلى برنامج table.sps، هناك برنامج صمم أيضاً بواسطة SPSS يقوم وحده بإنشاء ملفات التحليل و هو البرنامج CSPro.sps. يجب استخدام هذا البرنامج فقط عندما يتم تنفيذ جميع البرامج الأخرى لإنشاء ملفات التحليل و التأكد من أنها تعمل بصورة صحيحة فهو يفيد في إعادة إنشاء ملفات التحليل إذا تم تعديل أحد برامج إنشاء الملفات، فذلك يضمن أن جميع برامج إنشاء ملفات التحليل سيتم تنفيذها بالطريقة و الترتيب الصحيحين.

## أرشفة و توزيع البيانات

إن من الخطوات المهمة في معالجة البيانات – و يتم إغفالها غالباً – أرشفة وتوثيق ملفات البيانات. و سواء إذا كانت البيانات تنتشر على نطاق واسع أو ضمن مؤسسة واحدة فقط فإنه من الضروري وضع بعض الإرشادات فيما يتعلق بالتوزيع الجيد للبيانات مقدماً. تتطلب هذه الخطوات – أرشفة وتوثيق و توزيع البيانات - استثماراً للوقت والجهد و هي تستحق ذلك بالفعل لعدد من الأسباب، وهي:

- **زيادة فاعلية تكلفة جمع البيانات.** إن جمع بيانات المسوح هو نشاط مكلف و يحتاج إلى عدد كبير من الموارد البشرية. و لتبرير هذه التكلفة، يجب استغلال البيانات التي تم جمعها بأكثر قدر ممكن. ويزيد توفير ملفات البيانات لباحثين آخرين من فاعلية التكلفة المسوح.

● **تعزيز ملكية الدولة للبيانات وقبول نتائجها.** عندما يكون ملف البيانات متوفر للآخرين لاستخدامه، فإن عملية جمع البيانات تكتسب المصداقية، ويتم تصوّر جامعي للبيانات بأنهم على ثقة كبيرة بنتائج أعمالهم، وتعني إمكانية وصول ملف البيانات إلى باحثين آخرين بأن جميع النتائج يمكن نسخها والتحقق منها بواسطة آخرين.

● **القدرة على التحقق من التوجهات.** في الغالب تكون النتائج المنشورة عن المسوحات غير قابلة لعقد مقارنات. على سبيل المثال، يمكن أن يعرف تقرير لمسح ما البالغين على أنهم من هم بعمر الخامسة عشر عاماً فأكبر، بينما يمكن لتقرير مسح آخر أن يعرفهم على أنهم بعمر الثامنة عشر فأكبر. في هذه الحالة بدون ملفات البيانات، فإن أفضل ما يمكن عمله هو عقد مقارنة غير دقيقة لمجموعي النتائج. على كل حال، عندما تتوفر ملفات البيانات للمسح، غالباً ما يمكن إعادة جدولة النتائج حتى تكون قابلة للمقارنة المباشرة مما يسمح بتلخيص السياسات التي يجب وضعها.

● **القدرة على المقارنة بين الدول أو ضمن الدولة نفسها.** غالباً ما تكون مقارنة النتائج بين الدول عملية بناءة، سواء ضمن منطقة فرعية أو عبر المناطق. تسهّل هذه المقارنات معرفة المناطق التي يلزمها وضع برنامج محدد أو التي نفذت فيها برامج محددة بنجاح. كذلك، قد يكون من المفيد مقارنة النتائج من مسوحات مختلفة ضمن نفس الدولة. أحياناً يتم ذلك للتأكد من صحة نتائج غير متوقعة (مثلاً أن تكون نسبة وفيات الأطفال الرضع أقل مما هو متوقع) أو لتأييد منهجية معينة لجمع البيانات (مثل: الأخذ ببطاقات التطعيم بدلاً من تقرير الأمهات). ولإجراء هذه الأنواع من التحليلات، يلزم للباحثين الوصول إلى ملفات البيانات حتى يتم حساب الأرقام القابلة للمقارنة المباشرة.

● **تسمح بالتحليل المعمق لمواضيع محددة مهمة من قبل المتخصصين.** بسبب الضغوط للتقرير عن النتائج بزمن قصير وبسرعة، فإن المعلومات المقدمة في تقرير مسح ما عادة ما تتضمن النتائج الرئيسية للمسح فقط. وتسمح ملفات البيانات الموثقة بشكل جيد بإجراء تحليل معمق لمواضيع محددة، ويمكن القيام بمثل هذه التحليلات من قبل متخصصين قد لا يكونوا ضمن كادر المؤسسة التي قامت بجمع المعلومات.

يجب أرشفة ملف التحليل للمسوحات العنقودية متعددة المؤشرات وتوثيقه وتوزيعه. يجب أن تتضمن الوثائق المصاحبة لملفات التحليل على الأقل نسخة من التقرير الرئيسي ونسخة من الاستبيانات ووصفاً لتصميم العينة. ويجب أن تتضمن الوثائق أيضاً كتيب يتضمن أماكن ووصف كل متغير في ملف التحليل (يمكن إنشاء ذلك بسهولة في حزمة التحليل الإحصائي SPSS). يجب كذلك أرشفة نسخ من كل البرامج والملفات التي تم استخدامها خلال في المسح و تقديمها حين الطلب. يجب أيضاً إرسال نسخة من ملفات التحليل و وثائقها إلى المكتب الإقليمي لليونسف وإلى مكتب اليونيسف في نيويورك (قسم الإحصاءات والمراقبة)، وأخيراً يجب رسم سياسة وإجراء لتوزيع ملف البيانات إلى الآخرين.

## قائمة التحقق

### قبل القيام بالعمل الميداني:

- الحصول على أجهزة حاسوب والمعدات الأخرى لمعالجة البيانات.
- تجهيز غرفة أو مكان لمعالجة البيانات.
- توظيف مشرف معالجة بيانات وأفراد آخرين.
- وضع نظام لخطوات معالجة البيانات.
- تعديل البرامج لتتوافق مع الاستبيانات بمعالجة بيانات استبيان للاختبار.
- إدخال وتدقيق استبيانات الاختبارات القبلية.
- إنهاء البرامج بناء على الاختبارات القبلية والاستبيان النهائي.

### أثناء العمل الميداني:

- استلام الاستبيانات من الميدان.
- تعيين مدخل للبيانات الرئيسية.
- التحقق من بنية ملف البيانات الرئيسي.
- تعيين مدخل لبيانات التحقق.
- التحقق من أن ملفات البيانات الرئيسية وملفات بيانات التحقق متطابقة.
- حفظ نسخة احتياطية من ملف البيانات الخام.
- إجراء عملية التدقيق الثانوي.
- حفظ نسخة احتياطية من ملف البيانات النهائي.

### بعد العمل الميداني:

- تحويل البيانات إلى حزمة التحليل الإحصائي SPSS.
- إعادة ترميز المتغيرات.
- حساب وإضافة أوزان العينة ومعامل الثروة و البيانات الجغرافية GPS.
- تنفيذ برامج إنشاء الجداول.
- أرشفة البيانات و وضع سياسة لتوزيع البيانات و نظام أيضا(مثل: موقع إلكتروني).
- إرسال ملفات التحليل ووثائقها وجميع البرامج إلى اليونيسف.



