

### القسم الثالث: أخطاء الشمول/التغطية Coverage Errors

خلال هذا القسم سيتم تناول أحد أنواع أخطاء غير المعاينة، وهي أخطاء الشمول أو التغطية. وتحدث تلك الأخطاء في حالة عدم تمثيل الإطار أو السجل المستخدم في اختيار العينة لجميع مفردات المجتمع محل الدراسة، حيث قد ينتج عن ذلك تقديرات متحيزة بسبب العينة غير المثلة. فعلى سبيل المثال قد يتم استخدام دليل التليفونات في سحب عينة من مجتمع ما، وهذا الدليل لا يعتبر إطاراً ممثلاً للمجتمع، حيث لا يحتوي على أسماء الأشخاص الذين لا يمتلكون تليفونات، وبالتالي قد تنتج تقديرات متحيزة بسبب استخدام إطار غير ممثل للمجتمع محل الدراسة. فتنشأ أخطاء الشمول/التغطية عندما تكون بعض وحدات المجتمع محل الدراسة إما مستثناة أو مكررة بالقائمة المعدة كإطار يضم جميع وحدات المجتمع محل الدراسة.

#### ١.٣ مؤشر لحجم خطأ الشمول - نسبة الشمول (*Coverage Ratio*):

وتعرف نسبة الشمول بأنها نسبة المفردات المتضمنة في الإطار المستخدم -والتي تصف حجم شمولية الإطار- للمجتمع محل الدراسة. وتعتبر نسبة الشمول مؤشراً لحجم التحيز المحتمل في التقديرات دون التعرض لخصائص المفردات التي يشملها الإطار، أو خصائص المفردات التي لا يشملها الإطار. وإذا لم تكن هناك فروق في الخصائص بين المفردات التي يشملها الإطار والمفردات التي لا يشملها فإن عدم شمول الإطار لكل المفردات لن يؤدي إلى تقديرات متحيزة.

- النقطة (٣-١-١): إذا توافرت معلومات عن عدد المفردات التي تتضمنها الإطار المسحوب منه العينة، وعدد المفردات الكلية بمجتمع الدراسة يتم حساب مؤشر نسبة الشمول (عدد مفردات الإطار\عدد مفردات المجتمع). وتُضَمَّن نسبة الشمول -المُتَحَصَّل عليها- في الجدول بالملحق (أ).

### ٢.٣ مصادر أخطاء الشمول/التغطية

#### ١. مصادر أخطاء الشمول/التغطية التي تؤدي إلى تقديرات متحيزة:

هي تلك المصادر التي تؤدي إلى انحراف المعلومة المكونة أو الناتجة من استخدام العينات عن المعلومة الكاملة المحققة من المجتمع محل الدراسة. ويوجد نوعان من هذه المصادر وهي الشمول الجزئي (Under-coverage)، بالإضافة إلى تكرار التسجيل (Duplicate Listing).

#### أ- الشمول الجزئي (Under-coverage)

يحدث عدم الشمول نتيجة وجود عناصر في المجتمع المستهدف غير ممثلة بوحدات في الإطار المستخدم في سحب العينة الخاصة بالمسح. فعلى سبيل المثال يعد إطار بيانات عدادات الكهرباء إطاراً غير شامل، حيث أن ذلك الإطار لا يتضمن الأسر التي تقطن مساكن غير مزودة بالكهرباء، لذلك فإن استخدام هذا الإطار في المسوح الخاصة بالأسر التي تعيش في مساكن غير مزودة بالكهرباء يُعتَبَر مصدر من مصادر أخطاء الشمول. وكمثال آخر على عدم الشمول، استخدام إطار المؤسسات الحكومية والخاصة للتعرف على حجم ونوعية العمالة في المسح الخاص بالقوى العاملة، حيث أن ذلك الإطار لا يشمل القطاع غير الرسمي من الاقتصاد.

وقد تؤدي العناصر المفقودة التي لا يتضمنها الإطار غير الشامل لتقديرات متحيزة. وهذا النوع من أخطاء الشمول من الصعب قياسه لأنه بطبيعة تعريف هذا النوع من الأخطاء لا توجد أي معلومات متاحة عن المفردات المفقودة من الإطار، لكن على الرغم من ذلك يمكن قياس هذا النوع من الأخطاء في دراسات مخصصة لقياس جودة الأطر عن طريق المقارنة بالمصادر الخارجية للبيانات. مثل مقارنة الإطار المستخدم مع بيانات وسجلات آخر وأقرب تعداد تم إجراؤه.

وخلال تلك المناقشة يمكننا تسجيل نقاط الجودة التالية والتي يجب أن تؤخذ في الاعتبار عند

التقييم:

- النقطة (١-٢-٣): هل تم الإفصاح عن الإطار المستخدم؟
- النقطة (٢-٢-٣): هل يُمثِل هذا الإطار المجتمع دون أن يتغاضى عن أية جزئية به؟

### ب- التسجيل المكرر (Duplicate Listing)

يحدث التسجيل المكرر عندما يتم ربط مفردة واحدة في المجتمع المستهدف بأكثر من وحدة واحدة في الإطار المستخدم. وإذا لم تتم إزالة التكرارات من الإطار الذي يعاني من التسجيل المكرر قبل سحب العينة فإنه قد ينتج عن ذلك تحيز في التقديرات الناتجة. ولتجنب هذا التحيز فإنه يمكن وزن بيانات العينة لأخذ عدد الوحدات الموجودة في الإطار والتي تم ربطها بمفردات المجتمع المستهدف في الاعتبار.

وبذلك يمكننا استخدام نقاط الجودة التالية للتعبير عن هذه الجزئية:

- النقطة (٣-٢-٣): هل يمكن أن يحتوى الإطار على وحدات مُكررة؟، وبالتالي لا بد من حساب النسبة ١٠٠٪-نسبة احتواء الإطار على وحدات مكررة .
- النقطة (٣-٢-٤): هل توجد مفردات مُكررة بالعينة؟، وبالتالي لا بد من حساب النسبة ١٠٠٪-نسبة وجود مفردات مكررة بالعينة .

### ٢. مصادر أخطاء الشمول/ التغطية التي تؤدي إلى زيادة في تباين التقديرات:

وهي نوعان، الأول منها ينشأ نتيجة وجود عناصر في الإطار المستخدم في سحب العينة الخاصة بالمسح لا تنتمي للمجتمع محل الدراسة، والثاني منها ينشأ نتيجة وجود أخطاء في المعلومات الإضافية المساعدة.

أ- وجود عناصر في الإطار المستخدم في سحب العينة الخاصة بالمسح لا تنتمي

#### للمجتمع محل الدراسة

يمكن التعرف على المفردات غير المنتمية للمجتمع المستهدف واستبعادها خلال مسار عملية المسح، إلا أنه في الوقت نفسه تكون العينة الناتجة في النهاية أصغر مما كانت من المفترض أن تكون عليه مما يتسبب في زيادة تباين التقديرات الناتجة من المسح. لذلك تكون الإجابة على السؤال التالي محكاً أساسياً في الحكم على هذه الجزئية:

- النقطة (٣-٢-٥): هل توجد مفردات بالعينة لا تنتمي للمجتمع محل الدراسة؟، وبالتالي لا بد من حساب النسبة ١٠٠٪- نسبة وجود مفردات لا تنتمي للمجتمع محل الدراسة.

#### ب- وجود أخطاء في المعلومات الإضافية المساعدة

تستخدم المعلومات المساعدة الخاصة بالإطار المستخدم بشكل متكرر في كل من تصميم العينة لتقسيمها إلى طبقات تمثل المجتمع، وكذلك في حساب التقديرات للحصول على تقديرات دقيقة بأقل تكلفة ممكنة. وإذا كانت هذه المعلومات المساعدة غير كاملة أو تتضمن أخطاء فإن الفائدة التي تعود على التقديرات من ناحية زيادة الدقة ستتقلص، وسيزداد تباين التقديرات. ولسوء الحظ، لا يمكن قياس تلك النقطة بالشكل المباشر، إذ يتطلب الحكم عليها معرفة معلومات دقيقة جداً عن مجتمع الدراسة وهو ما لا يتوفر في معظم الأمور.

#### ٣. القوائم منتهية التأريخ أو غير الدقيقة:

يؤثر استخدام القوائم غير المحدثة كأطر لسحب العينات تأثيراً كبيراً على جودة البيانات. فعلى سبيل المثال تُعد الخرائط التي يتم تجهيزها للأعمال التعدادية إطاراً جيداً، ولكن في حالة عدم تحديثها باستمرار لمواكبة التوسعات العمرانية فإنها قد تؤثر بالسلب على جودة البيانات والتقديرات الناتجة من المسح.

بصفة عامة توجد ٣ طرق يمكن عن طريقها معرفة أن القائمة غير محدثة أو منتهية التأريخ:

- ١- عدم وجود عناصر جديدة في القائمة مما يسبب عدم الشمول، وإذا كانت هذه العناصر مختلفة في خصائصها عن العناصر الموجودة بالفعل في القائمة فإن ذلك قد يتسبب في تحيز التقديرات.
- ٢- عدم حذف عناصر قديمة من القائمة. وإذا كان من الممكن تحديد أن العنصر لا يجب أن تتضمنه القائمة فإن التأثير الوحيد على جودة البيانات سيكون زيادة في تباين التقديرات الناتجة نتيجة تقلص حجم العينة. أما إذا لم يكن من الممكن التمييز بين العناصر القديمة التي لا يمكن تتبعها والتقصي عنها وبين غير المجيبين- مثل حالة المسح البريدي لمؤسسات أعمال - فإن معدل عدم الاستجابة سيزداد.

٣-عدم وجود معلومات مساعدة حديثة. وتؤثر المعلومات المساعدة غير الدقيقة على جودة البيانات بطريقتين، فإذا كانت المعلومات مطلوبة لتتبع العناصر فإنه قد يكون من غير الممكن التفرقة بين العناصر التي لا يمكن تتبعها والتقصي عنها وبين العناصر غير المجيبة، مما يؤدي إلى تقليص حجم معدل الاستجابة الخاص بالمسح. أما إذا كانت المعلومات مطلوبة لتقسيم العينة إلى طبقات ممثلة للمجتمع أو حساب التقديرات فإن عدم وجود معلومات مساعدة محدثة يؤدي إلى زيادة تباين التقديرات.

ولدراسة جودة هذا المصدر من الأخطاء، علينا أن نجيب على السؤال التالي:

- النقطة (٣-٢-٦): إلى أي مدى تُعتبر القائمة -المستخدمة لسحب العينة- حديثة؟.