



الفصل الثالث:

جودة البيانات الإحصائية

د. لحسن عبد الله باشيوة

متخصص في مراقبة الجودة والإحصاء الرياضي

drbachioua@gmail.com

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

الأهداف التعليمية للمحاضرة

من المتوقع أن يصبح المتدرب عند نهاية هذا الفصل قادراً على القيام بما يلي:

✓ معرفة نوعية البيانات الإحصائية وتصنيفها

✓ استكشاف آليات تصنيف جودة البيانات

✓ معرفة أهمية جودة البيانات

✓ آليات تحسين وفحص البيانات الإحصائية

✓ مراقبة الجودة الإحصائية وإجراءات ضبط الجودة

✓ أنظمة وأدوات جودة البيانات الفعالة

أشارت المصادر العلمية التي تحدثت حول موضوع جودة البيانات إلى أهمية الموضوع، لما يترتب عليه من استخدام البيانات الإحصائية في مجالات مختلفة منها التخطيط والتنمية ورسم السياسات واتخاذ القرارات. هناك معايير وأبعاد للجودة لا بد من أخذها بعين الاعتبار في عملية إنتاج الرقم الإحصائي.

الجودة بشكل عام هي مقياس يقوم على أساس تفضيل سلعة أو خدمة معينة بالمقارنة مع غيرها من السلع والخدمات.

من الناحية الإحصائية: فإن السلعة هي (الرقم الإحصائي)، وبالتالي فإن جودة البيانات هي مدى تلبية الإحصاءات لحاجات المستخدمين واستجابتها لتوقعاتهم من حيث:

✓ المحتوى

✓ الشكل

✓ طريقة العرض

✓ الدلالات المعرفية التي تقدمها

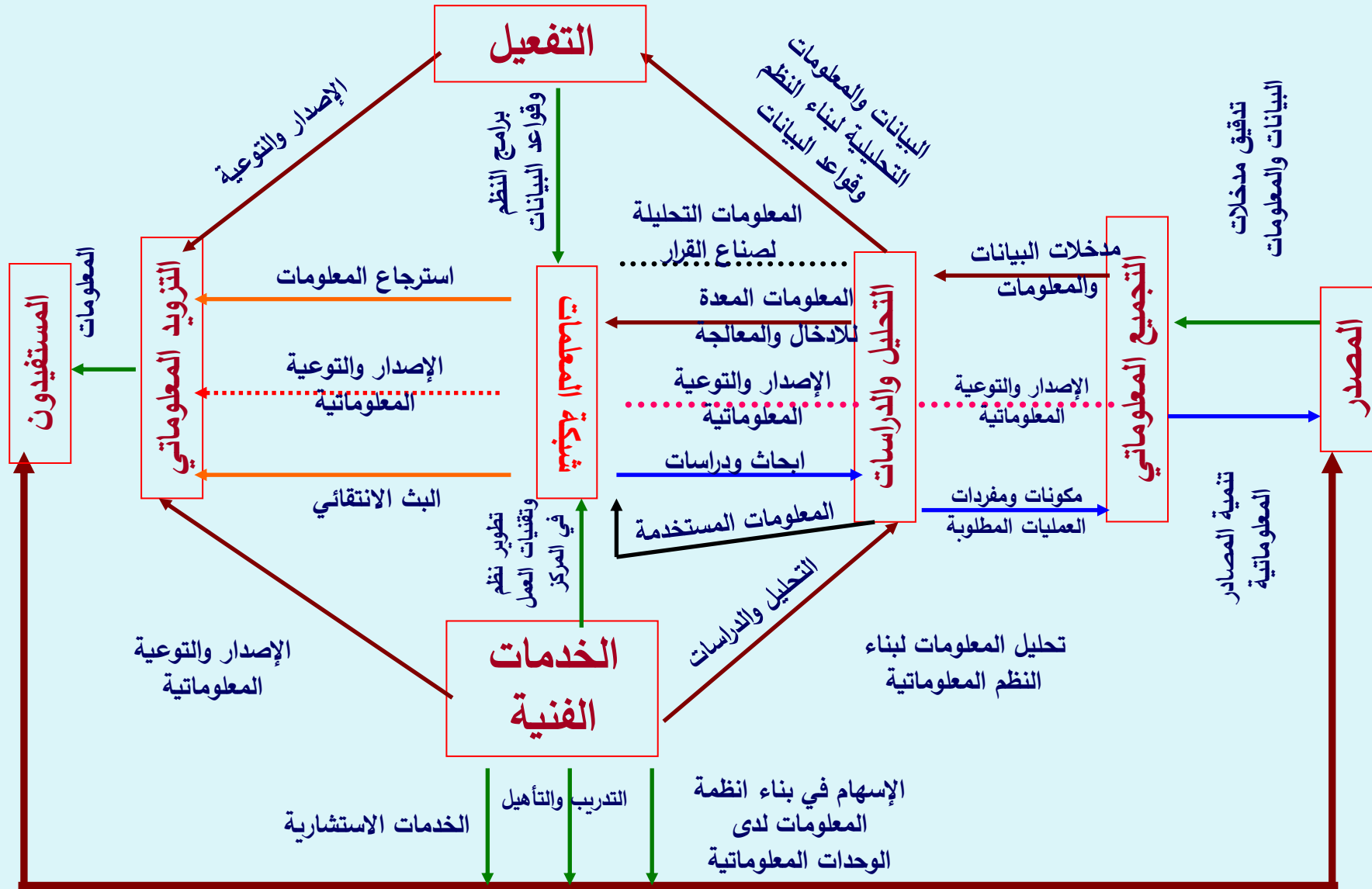
طرق جمع البيانات

تقوم الإدارة العامة لإحصاءات البيئة بتجميع و تحليل و تقويم البيانات الإحصائية من كل الوزارات و الهيئات المعنية في هذا المجال وذلك إما عن طريق **المقابلة الشخصية المباشرة** أو عن طريق **الإيميل** أو عن طريق **الفاكسات الرسمية**.

يتم عمل اجتماع تنسيقي ربع سنوي بين منسق الجهاز لإحصاءات البيئة و الجهات التي يتم الحصول منها على البيانات و ذلك لتذليل العقبات في عملية تدفق البيانات لإصدار النشرات و الدوريات في مواعيدها المحددة و بالجودة المطلوبة .

النظام الإحصائي وأنظمة الجودة

يشمل النظام الإحصائي في معظم الدول المؤسسة الإحصائية المركزية والوحدات الإحصائية الموجودة في المؤسسات الحكومية والبنك المركزي. وتشكل المؤسسة الإحصائية المركزية الركن المحوري في النظام الإحصائي الرسمي، حيث أنها تكون في العادة الجهة الوحيدة المخولة بإصدار الإحصاءات الرسمية، إذ تتولى جمع ومعالجة ونشر الإحصاءات الرسمية من مصادرها الأولية (المسوح والتعدادات) والثانوية (السجلات الإدارية المتوفرة لدى الوزارات والمؤسسات العامة). وبالتالي فان المؤسسة الإحصائية تشكل النقطة المحورية التي يتم فيها معالجة ونشر معظم الإحصاءات الرسمية.



مكونات النظام الإحصائي

يمثل النظام الإحصائي الجيد النقطة المحورية في إنتاج الإحصاءات الرسمية، وذلك لأنه يعتمد التوصيات الدولية كإطار نظري في المفاهيم والمصطلحات والمعايير والقياسات في مختلف المجالات الإحصائية. في العادة يغطي النظام الإحصائي أربعة برامج رئيسية هي الإحصاءات السكانية، والإحصاءات الاجتماعية، والإحصاءات الاقتصادية، والإحصاءات الجغرافية، ويتفرع من هذه التقسيمات:

الإحصاءات الاجتماعية والسكانية: تعتبر أساسا للتخطيط الاقتصادي والاجتماعي والبنية التحتية، وتشمل:

إحصاءات العمل والبطالة	إحصاءات الصحة
إحصاءات الفقر	إحصاءات التعليم والثقافة
إحصاءات النوع الاجتماعي ومشاركة المرأة	إحصاءات ظروف المعيشة
إحصاءات السكان	إحصاءات الأمن والعدالة
إحصاءات الطفل	إحصاءات أخرى

الإحصاءات الاقتصادية: تشمل الإحصاءات الاقتصادية على المواضيع التالية:

إحصاءات الصناعة

الحسابات القومية

إحصاءات الإنشاءات

إحصاءات التجارة

إحصاءات الأسعار

الإحصاءات الجغرافية: توفر البيانات الإحصائية ضمن الإحصاءات الجغرافية التالية:

إحصاءات الزراعة

إحصاءات البيئة

إحصاءات السياحة

إحصاءات المياه والطاقة والموارد الطبيعية

إحصاءات المساكن وظروف السكن

إحصاءات استعمالات الأراضي

إحصاءات النقل .

إحصاءات أخرى غير مصنفة ضمن البرامج الإحصائية الرئيسية: ظهرت مؤخرا مجالات جديدة يتم العمل

عليها حاليا ويعكف خبراء الإحصاء على تطوير قوائم مؤشرات خاصة بها، ومن هذه المواضيع:

إحصاءات حقوق الإنسان

إحصاءات الحكم الرشيد

إحصاءات القطاع غير المنظم

إحصاءات العنف الأسري

إحصاءات تكنولوجيا المعلومات والاتصالات .

النظام الإحصائي المركزي

يعتبر وجود نظام إحصائي مركزي ضروريا للأسباب التالية:

توفير الرقم الإحصائي الرسمي المبني على أسس علمية دقيقة وفقا لأحدث التوصيات والمعايير الدولية في مجال إعداد الإحصاءات الرسمية في جميع المجالات لتلبية احتياجات رسمي السياسات ومنتخذي القرارات والمهتمين وذلك من مصدر واحد بناء مؤسسة إحصائية، تلتزم بالعمل على تحقيق المتطلبات الإحصائية اللازمة لخطط التنمية، وتراعي الاستفادة من التجربة العالمية، وتوصيات الأمم المتحدة فيما يختص ببناء الأنظمة الإحصائية.

إنشاء نظام إحصائي شامل وموحد يكون بمثابة أداة لتحسين أداء الوزارات والمؤسسات في مجال عملها.

الحاجة لنشر الوعي الإحصائي عن طريق تقديم المعلومات من خلال وسائل الإعلام والتعاون مع الجامعات ومؤسسات البحث الأخرى وضمان حق المواطن في الوصول إلى المعلومات على أساس أن الإحصاء سلعة عامة.

المنهجيات المتبعة في مجال العمل الاحصائي

- التنوع في إصدار البيانات الإحصائية.
- الشفافية و الوضوح في عرض البيانات.
- إظهار جهود الدولة في الحد من التلوث.
- تبادل البيانات و المعرفة بين الهيئات و المؤسسات المعنية بالبيئة
- تحليل البيانات من المصادر و تكوين سلاسل زمنية لسهولة المقارنة بين تلك البيانات و توضيح مدى الاختلاف الحاصل بالسلب أو بالايجاب .
- نشر البيانات بصورة تتسم بالوضوح و الشفافية و بحالة ترضى مستخدم البيان و متخذ القرار.
- الآخذ في الاعتبار المعايير الإحصائية العالمية و استخدام المفاهيم و التصانيف و الأساليب الدولية مما يعزز المقارنات الدولية.

جودة البيانات

فهل للبيانات جودة أو هل يوجد شيء اسمه جودة البيانات؟

تعريف البيانات **Data** : هي العناصر التي نستخلص منها المعلومات بعد

المعالجة (ولا تكون مفيدة بمفردها) وتكون البيانات هي المادة الخام

للمعلومات ، وتنقسم البيانات إلى النوعين الأساسيين التاليين:

1. بيانات رقمية : وهي البيانات التي تأخذ قيمة محددة لا تخرج عنها ، ومثال ذلك حروف الهجاء.

2. بيانات تمثيلية : وهي البيانات التي تأخذ قيمة عديدة غير محدودة ، نحو شدة الصوت أو درجة الحرارة.

مراجعة البيانات المتاحة لدينا و مقارنتها بالمعدلات العالمية مع الأخذ في

الاعتبار الحدود الدولية العظمى و الصغرى للمعايير البيئية.

التعاون المستمر مع مراكز المعلومات بالوزارات و الهيئات المعنية لضمان

جودة البيانات.

وبناء على هذا التعريف البسيط نصل إلى معنى هام هو أنه إذا كانت المادة الخام للمعلومات غير جيدة فسوف نحصل بالتأكيد على معلومات غير جيدة وسنبنى عليها قرارات سوف تكون باحتمال كبير خاطئة. وهنا نصل إلى ضرورة التأكيد على أن تكون البيانات جيدة وتكون البيانات جيدة إذا ما كانت:

1. دقيقة ،

2. مكتملة،

3. متسقة (أي ذات صلة بموضوعها).

4. مرتبطة ارتباطا صحيحة بزمنها

5 . موثوقا فيها .

6. يمكن فهمها .

7. يسهل التعامل معها لاستخلاص معلومات منها.

إذا نصل الآن لإجابة عن تساؤلنا السابق وهو أنه للبيانات جودة لأن البيانات مثلها مثل أي شئ اعتباري لها خصائص وكل شئ اعتباري له خصائص تكون له جودة وهنا نعيد الطرح من جديد إذا كانت للبيانات جودة فما هو السبيل لتحقيق هذه الجودة للبيانات. إن النظام الذي تسقى منه البيانات عليهم أن يراعون كيف يزرعون خصائص جودة البيانات في كل النماذج والتقارير التي يبنونها في نظمهم الإدارية .

● تحسين جودة البيانات

- إرساء جسور متطابقة وفهم مشترك بين جامعي البيانات والمختصين في دائرة الاحصاء والمعلومات .
- اعتمادا على البيانات الإحصائية لبناء الخطط المستقبلية باعتبارها حجر الأساس .
- تقييم العينات الإحصائية المجمعة
- تقييم ومراجعة المنهجيات الإحصائية المستخدمة وأسلوب جمع البيانات وآلية تنفيذها والإشراف عليها .
- نشر المخرجات الإحصائية بمختلف الأساليب ذات الكفاءة .

أبعاد الجودة السبعة للبيانات

هناك مجموعة من الخواص والسمات في المنتج الإحصائي تجعله يلبي حاجات المستخدمين تسمى أبعاد الجودة السبعة وهي:

Relevance الصلة بالواقع ➤

Accuracy الدقة ➤

Timeliness and Punctuality الوقتية والانتظام ➤

Accessibility and clarity إمكانية الوصول والوضوح ➤

Comparability القابلية للمقارنة ➤

Coherence الاتساق ➤

Completeness الاكتمال ➤

التحقق من البيانات (Data Validation): هي واحدة من العمليات الموجودة و التي يقوم بها الباحث ليتأكد من أن البيانات المدخلة نظيفة ولا تهدف إلى تخريب البرنامج أو أي شيء من هذا القبيل .إن عملية التحقق من البيانات تستخدم بعض القواعد للتأكد من أن البيانات التي يتم إدخالها للبرنامج صحيحة ومفيدة وآمنة.

حفظ رقمي (Digital preservation): هي سلسلة من الإجراءات اللازمة لضمان استمرار إمكانية الوصول إلى المواد الرقمية كلما كان ذلك ضرورياً، فهو يجمع بين السياسات والاستراتيجيات والإجراءات لضمان الوصول وتهيئة وتوليد محتوى رقمي بغض النظر عن التحديات التي تواجهه فشل الوسائط والتغير التكنولوجي .هدف الحفظ الرقمي هو استخلاص دقيق لمحتوى موثوق على مر الزمن .

● **سلامة البيانات (Data integrity):** تشير إلى صيانة وضمان دقة البيانات عبر دورة حياتها، وهي الجانب الحاسم في تصميم وتنفيذ واستخدام أي نظام يخزن العمليات أو يقوم باسترداد البيانات ويمكن ان يكون لها معاني مختلفة على نطاق واسع اعتمادا على السياق المحدد حتى تحت نفس المنظور العام للحوسبة ، وتنص هذه المادة لمحة عامة عن بعض أنواعها المختلفة والمخاوف من سلامة البيانات . سلامة البيانات هي عكس تلف البيانات والذي هو شكل من أشكال فقدان البيانات القصد العام من أي تقنية .

● **تلف البيانات (Data corruption):** تشير إلى أخطاء في البيانات التي تحدث أثناء الكتابة أو القراءة أو التخزين أو النقل أو المعالجة والتي تُدخل بيانات غير مقصودة على البيانات الأصلية . أنظمة النقل والتخزين في الكمبيوتر لديها عدد من التدابير لضمان سلامة البيانات، و عدم وجود أخطاء فيها.

أهمية جودة البيانات

تأتي أهمية جودة البيانات نظراً لاستخدام الرقم الإحصائي في كل عمليات الرصد والتخطيط والتحسين والتقييم واتخاذ القرارات. يجب على القائم على جودة البيانات أن يراعي الأمور الآتية للحفاظ على جودة البيانات:

- أن يولي أولوية للجودة بغرض كسب ثقة المستخدمين.
- وصف جودة البيانات والمنهجية المستخدمة بصورة دقيقة، مع استخدام نظام مراقبة جودة فعال.
- عقد حوار المنتجين والمستخدمين والمقيمين.
- إعادة الفحص على البيانات وتعدينها.
- تعزيز مبدأ الجودة في العمليات الإحصائية المختلفة.
- تتبع التسلسل الهرمي لتدفق البيانات الإحصائية وتتبعها ومنع وقوع الأخطاء والحصول على البيانات الموثوقة من أول خطوة.

مراقبة الجودة الإحصائية وإجراءات ضبط الجودة

هي مسار منهجي تستطيع من خلاله الأجهزة الإحصائية جمع وتقييم عناصر إثبات موضوعية بخصوص الأخطاء الإحصائية الممكنة، وذلك بغرض البت في هذه الأخطاء بالرجوع للمعايير المتبعة في هذه الأجهزة. وهي عملية مراجعة منتظمة للأنشطة، وتقييم مستمر للأنشطة، وهي مقارنة سير العمل والإنتاجية مع الجدول الزمني.

يتم العمل على ضبط الجودة في المشاريع والمسوح الإحصائية باستخدام الوسائل الآتية:

قاعدة بيانات شروط ضبط إجراءات الجودة في المشاريع الإحصائية

التقييم الذاتي لمدرء المشاريع DESAP

تقارير ودراسات متخصصة بالجودة

فحص البيانات

تطبيق استمارة ميثاق الممارسات الأوروبية COP

النظام العام لنشر البيانات GDSS

النظام الخاص لنشر البيانات SDDS

جودة الإحصاءات

يعتبر مفهوم جودة الإحصاءات الرسمية مفهوما معقدا. ففي المناقشات الدولية الحديثة، تم توسيع المفهوم إلى ما وراء الاهتمامات التقليدية في دقة الأرقام ليشمل الجودة الكلية في نظر المستخدمين. وجاء هذا التركيز على رأي المستخدمين من خلال الحركة الداعية إلى منهاج "الإدارة الكلية للجودة" والذي تم تطويره وتطبيقه أولا في الإنتاج الصناعي للمنتجات الاستهلاكية، والآن أصبح يطبق في جميع المجالات بما فيها المؤسسات الإحصائية. وقامت بعض المؤسسات الإحصائية الوطنية بتبني فلسفة ومناهج "الإدارة الكلية للجودة" كنظام منهجي لتحسين الجودة.

يتطلب تطبيق نظام "الإدارة الكلية للجودة" في الإحصاء وضع اهتمامات المستخدمين في موقع مركزي، والعناية بجودة الإنتاج، واستخدام تقنيات العمل الجماعي وتوفير أدوات إحصائية مفيدة في التشخيص، والمراقبة والأمور الأخرى الإحصائية. وعلى الأجهزة الإحصائية، سواء تبنت "الإدارة الكلية للجودة" أو فلسفة إدارة جودة أخرى لتكون قاعدة لسياسة تحكم ومراقبة، أن تقرر أبعاد جودة الإنتاج الإحصائي.

الشروط الأساسية للجودة، وتشمل:

البيئة القانونية المؤسسية

الموارد للمؤسسة الإحصائية

الصلة بالموضوع

الموضوعية، وتشمل:

الكفاءة المهنية

الشفافية

المعايير الأخلاقية

المنهجية، وتشمل:

المفاهيم والتعريفات المستخدمة في العمل الإحصائي

الشمول

التصنيفات المستخدمة

الدقة والموثوقية، وتشمل:

1. مستوى الثقة بمصادر البيانات
 2. تقييم البيانات الأولية
 3. الأساليب الإحصائية المستخدمة في العمل الإحصائي
 4. تقييم البيانات الوسطية والمخرجات الإحصائية وإقرارها
 5. نماذج الافتراضات المستخدمة في التحليل
- إجراءات تعديل البيانات التي تتطلب ذلك والإفصاح عن الفرضيات المستخدمة في التعديل
الإعلام المسبق عن أي تعديل يطرأ على المنهجية المستخدمة في العمل
المنفعة، وتشمل العناصر التالية:

1. الدورية والحدثة
2. اتساق البيانات
3. سياسة وممارسات التعديل

سهولة الاطلاع، وتشمل:

1. سهولة الاطلاع على البيانات الإحصائية
2. سهولة الاطلاع على البيانات الوصفية
3. تقديم المساعدة للمستخدمين

قوة المعلومات

وجود جهاز للإحصاء والمعلومات ومدى استقلاليته ومهنيته وكفاءته .

توفر المعلومات بالجودة والمصداقية والوقت المناسب

وجود وتطبيق منهج وطني في مجال الإحصاء والمعلومات .

قدرة الدولة في إنتاج مؤشرات للقياس والتقييم في مجالات القوة الإستراتيجية الشاملة .

أوضاع الوعي الإحصائي و المعلوماتي .

توفر كوادر الإحصاء والمعلومات .

كفاءة التشريعات المتعلقة بالمعلومات والإحصاء .

توفر التكنولوجيا المساندة للمعلومات والإحصاء

{ نظم تقنية المعلومات وتكنولوجيا الاستشعار عن بعد إلخ }

ويشمل مفهوم الجودة للإحصاءات الرسمية وفقا لإطار تقييم جودة البيانات الإحصائية الصادر عن صندوق النقد الدولي (DQAF) المحاور التالية:

لقد قامت عدة منظمات دولية بتطوير أدلة لمراقبة التقدم المحرز في تطوير الإحصاءات ومن أهمها دليل تقييم جودة البيانات (DQAF)، وأدلة معياري النشر الصادرة عن صندوق النقد الدولي (المعيار العام لنشر البيانات (GDDS) والمعيار الخاص لنشر البيانات (SDDS)). وهذه الأدلة تصلح لأن تستخدمها الأجهزة الإحصائية في تقييم جودة الإحصاءات الصادرة عنها ومدى شمولها والتزامها بالمبادئ الأساسية للإحصاءات الرسمية الصادرة عن الأمم المتحدة.

إن انضمام الدول إلى معياري النشر (GDDS و SDDS) الصادرة عن صندوق النقد الدولي يعتبر مؤشرا قويا على مدى التقدم في توفير الإحصاءات الاقتصادية والاجتماعية، حيث تعتبر نماذج تقييم المؤسسات الإحصائية (نموذج C) ونماذج تقييم البيانات (نموذجي A و B) مناسبة للتقييم. نلاحظ أن نموذج A يتحدث عن شمولية النظام من حيث آليات جمع البيانات وطبيعة مصادر البيانات والوقتية وطرق التحليل المستخدمة والمفاهيم والمصطلحات والتصنيفات المستخدمة والمجالات التي تغطيها الإحصاءات. فيما يتناول نموذج B فئات البيانات والمؤشرات ونماذج النشر وخطط تطوير الإحصاءات والجودة وخصائص البيانات. أما نموذج C فيتناول التكامل بين البيانات التي تقوم المؤسسة الإحصائية بإعدادها والإجراءات المتخذة لتوفيرها للعموم.

تطبيق استمارة ميثاق الممارسات الأوروبية COP

ميثاق الممارسات Code of Practice: مجموعة من المبادئ التوجيهية واللوائح التي يجب إتباعها من قبل أعضاء مهنة معينة.

ميثاق ممارسات الإحصاءات الأوروبية: قواعد تضع معايير لتطوير وإنتاج ونشر الإحصاءات الأوروبية، وقد تم بناء هذه القواعد استناداً إلى التعريفات المتعلقة بجودة الإحصاءات تستهدف هذه القواعد جميع مساحات العمل ذات العلاقة من البيئة المؤسسية، وعمليات الإنتاج الإحصائي وصولاً إلى المخرج النهائي: الإحصاءات الرسمية الأوروبية.

لقد اعد ميثاق الممارسات لخدمة غرضين أساسيين:

✓ زيادة ثقة الجهات الخارجية في المؤسسة الإحصائية

✓ تحسين جودة البيانات في المؤسسة الإحصائية

تم تصميم استمارة قواعد الممارسات لتتوافق مع الإطار الأوروبي للجودة والذي يعتبر أكثر تحديداً وتفصيلاً من المبادئ الأساسية للإحصاءات الرسمية الصادرة عن الأمم المتحدة، وقد تم تطويره تماشياً مع أنظمة الجودة المتبعة من قبل يوروستات وأجهزة إحصائية أخرى.

تطبيق استمارة ميثاق الممارسات الأوروبية COP

تتكون استمارة التقييم الذاتي من 15 مبدأ تغطي ثلاثة مواضيع رئيسية وهي:

✓ البيئة المؤسسية (الجودة الهيكلية) Structural Quality

✓ العمليات الإحصائية (جودة العملية) Process Quality

✓ المخرجات الإحصائية (جودة المخرجات) Product Quality

مراجعة النظراء Peer Review: تعرف مراجعة النظراء على أنها العملية

المستخدمة لفحص العمل الذي تؤديه جهة معينة للتأكد من أنها تلي

معايير محددة، ويقوم بهذا الفحص جهة خارجية مؤهلة ومحايدة، ويعتقد أن

مراجعة النظراء ضرورية لتحديد الأخطاء بسرعة وسهولة، وتسريع الوقت

الذي تستغرقه عملية تشخيص الأخطاء وتصحيحها.

يحتوي ميثاق الممارسات للإحصاءات الرسمية على مجموعة من المبادئ الأساسية التي تحدد الممارسات الفضلى في العمل الإحصائي الرسمي، وهذه المبادئ هي:

1. الصلة بالموضوع
2. استقلالية الإحصاءات
3. الجودة
4. إتاحة البيانات للاستخدام العام
5. حماية سرية البيانات
6. التوازن بين احتياجات المستخدمين والعبء على المبحوثين
7. التكامل وتراكم الخبرة والإبداع
8. تكلفة إنتاج الإحصاءات الرسمية وجدواها .

شكرا لسلامة

جزاكم الله خيرا على حسن المتابعة

هل من أسئلة .