



توصیه‌های بین‌المللی در به‌کارگیری داده‌های ثبتی و اداری در سرشماری نفوس و مسکن

ژویشکده‌ی آمار

مجری: پریا ترابی

همکار: لیدا کلهری

زمستان ۱۴۰۴

اهداف کلی

بخش اول: کلیات

- ✓ نگاهی به انواع رویکردهای سرشماری‌های جمعیت در جهان
- ✓ پیش‌شرط‌های تعیین‌شده در حرکت به سمت رویکرد ثبتی مبنا
- ✓ اهم چالش‌ها و مخاطره‌های حرکت به سمت سرشماری ثبتی مبنا
- ✓ روش‌های مختلف به‌کارگیری داده‌های اداری در سرشماری
- ✓ تعاریف و مفاهیم مهم از جمله

- انواع ثبت‌های پایه و تخصصی مورد استفاده در سرشماری
- راهکارها در صورت عدم دسترسی به داده‌های منابع اداری
- تبدیل داده‌های اداری به پایگاه‌داده‌ی آماری و ایجاد ثبت آماری جمعیت (SPR) و تفاوت آن با ثبت جمعیت

بخش دوم: ارزیابی کیفیت

- ✓ کیفیت منبع
- ✓ کیفیت داده‌های ورودی
- ✓ کیفیت فرایند
- ✓ کیفیت برون‌داد

بخش سوم: مروری بر مطالعات تطبیقی

- ✓ استرالیا
- ✓ اندونزی
- ✓ جمهوری کره
- ✓ ترکیه

بخش اول: کلیات



سرشماری‌های جمعیت در جهان

نتایج آمارگیری کوتاه ۱۱ سؤالی از ۱۵۸ کشور با هدف گردآوری اطلاعات در مورد برنامه‌ها و نتایج کشورها در سرشماری سال ۲۰۲۰

جدول ۱- روش‌های اصلی سرشماری‌های نفوس و مسکن سال ۲۰۲۰

کشورهایی که در آمارگیری شرکت کرده‌اند	تعداد و درصد از ۱۵۸ کشور پاسخ‌گو	آفریقا (۳۸)	آسیا (۳۶)	اروپا (۴۱)	آمریکای لاتین و منطقه‌ی کارائیب (۲۸)	آمریکای شمالی (۴)	اقیانوسیه (۱۱)	انواع سرشماری
	۱۱۲ (۷۱ درصد)	۳۷	۲۵	۱۲	۲۸	۱	۹	سرشماری سنتی
	۱۶ (۱۰ درصد)	۰	۳	۱۲	۰	۱	۰	سرشماری ثبتی مبنای کامل
	۲۱ (۱۳ درصد)	۱	۵	۱۱	۰	۲	۲	سرشماری ترکیبی (ثبت‌ها و شمارش میدانی کامل)
	۸ (۵ درصد)	۰	۳	۵	۰	۰	۰	سرشماری ترکیبی (ثبت‌ها و آمارگیری نمونه‌ای)
	۱ (۱ درصد)	۰	۰	۱	۰	۰	۰	سرشماری غلتان

منبع: (UNSD, ۲۰۲۰)

جدول ۲- نوع سرشماری نفوس و مسکن سال‌های ۲۰۰۰ تا ۲۰۳۰ کشورهای منطقه‌ی UNECE

انواع سرشماری	سال ۲۰۰۰	سال ۲۰۱۰	سال ۲۰۲۰	سال ۲۰۳۰
سرشماری سنتی*	۴۰	۳۵	۱۴	۲
سرشماری ترکیبی	۵	۱۰	۱۷	۱۴
سرشماری ثبتی مبنای	۳	۹	۱۳	۲۳

*این رده شامل سرشماری غلتان در فرانسه است.

۱ • چشم‌انداز راهبردی بلندمدت NSO برای نظام آماری ثبتی مبنا (بدون در نظر گرفتن پیش‌شرط‌های لازم به‌عنوان پایه و اساس نظام ثبتی مبنا، حرکت به سمت سرشماری ثبتی مبنا در کوتاه‌مدت امکان‌پذیر نیست).

۲ • حمایت سیاسی و چارچوب قانونی (شامل حمایت همه‌جانبه و نیز تغییرات قانونی لازم و رفع موانعی است که ممکن است مانع یا بازدارنده‌ی متولیان ثبت از ارائه‌ی داده‌های مورد نیاز به NSO برای سرشماری یا هر فعالیت آماری دیگری شود).

۳ • حمایت عموم مردم، محرمانگی و مشارکت کاربران (استفاده از داده‌های اداری در سرشماری باید مطابق با ششمین اصل بنیادین آمار رسمی باشد).

۴ • همکاری بین NSO و متولیان ثبت (علاوه بر چارچوب قانونی و ارتباط‌های مؤثر با سایر متولیان، امضای تفاهم‌نامه‌های همکاری می‌تواند به حمایت از برنامه‌های آماری NSO و به‌ویژه فرایند سرشماری کمک کند).

۵ • دانش ثبت‌های اداری مربوط و ارزیابی اولیه‌ی کیفیت آن‌ها (این گام شامل ارزیابی منابع بالقوه و مناسب برای تمام موضوعات سرشماری مورد نیاز کاربران اصلی و ثبت‌های پایه است).

۶ • نظام شناسایی یکپارچه (برای تسهیل اتصال داده‌های اداری از منابع مختلف که به واحد یکسانی مربوط می‌شوند، وجود شماره‌های شناسایی یکتا در سطح رکورد واحد که در تمام ثبت‌های مربوط مشترک باشند، لازم است).

۷ • زیرساخت و ظرفیت نهادی (ایجاد زیرساخت‌های لازم، طراحی و اجرای برنامه‌های آموزشی برای کارکنان NSO جهت کسب مهارت‌های آماری و فناوری اطلاعات لازم، امری ضروری است).

اهم چالش‌ها و مخاطره‌های حرکت به سمت سرشماری ثبتی مبنا

۱. وابستگی بیشتر NSO به متولیان دولتی،
۲. تفاوت در مفاهیم، تعاریف و نظام‌های رده‌بندی مورد استفاده در منابع داده‌های اداری در مقایسه با سرشماری‌ها،
۳. تضاد داده‌های یک ثبت به دلیل تفاوت‌ها در تعاریف، رده‌بندی‌ها، زمان‌بندی یا به دلیل خطا در منبع با داده‌های ثبت دیگر،
۴. تأخیر در تحویل داده‌های اداری و فراداده‌های مربوط به NSO‌ها و ایجاد مشکلاتی برای آن‌ها در مورد روزامدی آمارهای ثبتی مبنا،
۵. فقدان نظام شناسایی یکپارچه برای اتصال داده‌ها از ثبت‌های مختلف،
۶. وجود داده‌های گمشده و کم‌پوشانشی برخی از جمعیت‌ها در منابع اداری،
۷. حفظ زیرساخت‌های فناوری اطلاعات NSO،
۸. افزایش نگرانی‌ها در ارتباط با حریم خصوصی در جامعه،
۹. عدم اعتماد آمارشناسان در استفاده از منابع اداری و ترس از خطر تعدیل شغلی آن‌ها به دلیل استفاده از این داده‌ها، و
۱۰. کاهش علاقه‌مندی عمومی نسبت به سرشماری و نتایج آن.

روش‌های مختلف به کارگیری داده‌های اداری در سرشماری



ایجاد چارچوب‌های
نمونه‌گیری



پشتیبانی از عملیات
میدانی



ایجاد ثبت جمعیت



جایگزینی یا ایجاد
متغیرهای جدید
سرشماری



ارزیابی/بهبود کیفیت
داده‌های سرشماری

انواع ثبت‌های پایه‌ی مورد استفاده در سرشماری

ثبت جمعیت

شامل متغیرهای و ویژگی‌های فردی از قبیل نام، جنس، تاریخ و محل (یا کشور) تولد، وضع زناشویی، شماره‌ی شناسایی فردی (PIN)، تابعیت و نشانی اقامت‌گاه معمول و تاریخ ورود یا خروج از آن نشانی

ثبت املاک و مستغلات (سکونت‌گاه‌ها و ساختمان‌ها)

مانند موقعیت ساختمان که به‌طور معمول به‌عنوان نشانی رسمی شناخته می‌شود و ممکن است شامل مختصات جغرافیایی نیز باشد، نوع و هدف هر ساختمان به ویژه آن دسته از ساختمان‌هایی که برای اهداف مسکونی مورد استفاده قرار می‌گیرند، (خواه ساختمان‌های تصرف‌شده یا خالی باشند) یا غیرمسکونی، و ویژگی‌های ساختمان شامل جزئیات مربوط به قدمت، طرح، اندازه و تأسیسات آب، گاز و برق

ثبت بنگاه‌ها و کارگاه‌های کسب و کار

شامل نشانی از جمله کد جغرافیایی رسمی در پایین‌ترین سطح، بخش و زیر بخش نهادی که برای تمایز بین فعالیت‌های تجاری و غیرتجاری و بین بنگاه‌های دولتی و خصوصی استفاده می‌شود، فعالیت‌های اقتصادی اصلی و ثانویه که برای تخصیص صنعت مطابق با رده‌بندی استاندارد صنعتی در سطح ۳-رقم استفاده می‌شود، و اندازه‌ی نیروی کار برحسب تعداد افراد شاغل و کارکنان

ثبت نشانی

هیچ تعریف رسمی مورد توافق بین‌المللی از آن وجود ندارد و در واقع رابطه‌ی بین ثبت‌های ساختمان/سکونت‌گاه و نشانی مختص هر کشور است. هر کشور باید قادر به تهیه‌ی فهرستی از مکان‌هایی باشد که مردم در آن ساکن هستند، همچنین باید قادر به تعیین موقعیت و اتصال مکان به ثبت جمعیت باشد. در سرشماری‌های ثبتی مبنا یا ترکیبی، ثبت نشانی برای حصول اطمینان از هماهنگ‌سازی و کامل‌بودن پوشانش یاری رسان است

انواع ثبتهای تکمیلی (تخصصی یا موضوعی) مورد استفاده در سرشماری



عدم دسترسی به داده‌های منابع اداری

گردآوری تمام متغیرهای مورد نظر سرشماری از ثبت‌ها عملی و ممکن نیست. به‌طور مثال:

- (الف) اطلاعات مربوط به روابط داخلی بین اعضای خانوار، مانند همباشی بدون ازدواج (Cohabitation Outside Marriage) که برای به‌دست آوردن وضعیت ترکیب‌بندی خانوادگی و خانوار ضروری است، در هیچ ثبت اداری نگهداری نمی‌شود،
- (ب) به دلیل تعلق خاطر و حساسیت ویژگی‌های مرتبط با فرهنگ مانند قومیت، مذهب و زبان، تعداد کمی از ثبت‌های اداری (در صورت وجود) چنین داده‌هایی را در اختیار دارند،
- (پ) منابع اداری مربوط به فعالیت اقتصادی نیروی کار (مانند ثبت مشاغل یا کسب‌وکارها) اغلب دارای سطح مطلوبی از اطلاعات دقیق در مورد نوع و محل کار یا کارکنان بخش غیر رسمی نیستند تا NSOها بتوانند رده‌بندی شغل و صنعت را که مطابق با استانداردهای بین‌المللی است، ایجاد کنند.
- (ت) متغیرهای کلیدی مسکن مانند تعداد اتاق‌ها یا اتاق‌های خواب یا وجود بالابر یا دسترسی‌پذیری معلولان ممکن است در هیچ‌کدام از ثبت‌های اداری سکونت‌گاه ثبت نشوند.

در صورت وقوع شرایط مذکور، راهکارهای دسترسی یا منابع داده‌های جایگزین می‌تواند شامل استفاده از موارد زیر باشد:

- آمارگیری‌های نمونه‌ای موجود که توسط NSO اجرا شده است (مانند آمارگیری نیروی کار یا آمارگیری هزینه و درآمد خانوار)،
- آمارگیری پوشانش جمعیت یا سایر آمارگیری‌های جمعیت‌شناختی، یا
- منابع داده‌های تجاری از جمله مه‌داده‌ها.

❖ به‌عنوان مثال، هیچ ثبتی از ساختمان‌ها و سکونت‌گاه‌ها که کل آلمان را پوشش دهد، وجود ندارد و در نتیجه برون‌دادهای سرشماری مورد نیاز برای سرشماری مسکن در سال ۲۰۱۱ (علاوه بر برخی اطلاعات کمکی مربوط به ترکیب خانوار) از داده‌های گردآوری‌شده از طریق آمارگیری پستی به‌دست آمد.

❖ مثال دیگر، سرشماری مجازی هلند است که به آمارگیری‌های موجود برای گردآوری اطلاعات در مورد شغل و صنعت افراد متکی است زیرا ویژگی‌های اقتصادی در هیچ ثبت اداری در هلند موجود نیست.

تبدیل داده‌های اداری به پایگاه داده‌ی آماری و ایجاد ثبت آماری جمعیت (Statistical Population Register (SPR))

- SPR گردآوری شاخص محور و نظام مند رکوردهای فردی برای هر فرد ساکن (اعم از شهروندان ملی و اتباع خارجی) در کشور است.
- SPR چارچوب جمعیت پایه را برای طراحی و اجرای آمارگیری‌های آماری ارائه می‌دهد.
- SPR باید به صورت منظم بر اساس قوانین حاکم بر ثبت‌های اداری جمعیت با اطلاعات مربوط به‌روز شود.
- ایجاد SPR یک گام اساسی در جهت دستیابی به استفاده از داده‌های اداری برای اهداف سرشماری است.
- SPR متفاوت از ثبت افراد برای اهداف اداری است و برای اهداف آماری مورد استفاده قرار می‌گیرد.
- SPR به‌طور معمول با استفاده از ثبت اداری جمعیت، توسط NSO تولید می‌شود. اگر ثبت جمعیت متمرکز وجود نداشته باشد، می‌توان با یکپارچه‌سازی ثبت‌های محلی جمعیت و اتصال آن‌ها با سایر ثبت‌های موجود، ثبت آماری جمعیت را ایجاد کرد.

وجه تمایز ثبت جمعیت و ثبت آماری جمعیت (SPR)

ثبت آماری جمعیت

- هدف ارائه‌ی اطلاعات آماری درباره‌ی جمعیت است.
- از ثبت اداری جمعیت اقتباس شود.
- توسط اداره‌های آمار ملی (NSO) ایجاد و نگهداری می‌شود.

ثبت جمعیت

- هدف اصلی ثبت اطلاعات افراد برای اهداف اداری دولتی است.
- ثبت اداری جمعیت نیز نامیده می‌شود.
- توسط برخی از نهادها ایجاد و نگهداری می‌شود.

بخش دوم: ارزیابی کیفیت





چهار مرحله‌ی ارزیابی کیفیت داده‌های اداری

کیفیت منبع

ارزیابی کیفیت منابع داده‌های اداری
مورد استفاده در سرشماری (ارزیابی
فراداده-مبنا)

کیفیت داده‌های ورودی

ارزیابی کیفیت داده‌های اداری خام
ارائه‌شده توسط سازمان‌های اداری به
NSO ها

کیفیت فرایند

ارزیابی تغییرات در کیفیت داده‌های
حاصل از یکپارچه‌سازی و پردازش
داده‌های اداری

کیفیت برون‌داد

ارزیابی کیفیت کلی نتایج آماری
منتشرشده برای کاربران

کیفیت منبع – ارزیابی مرتبط بودن

کیفیت منبع

مرتبط بودن

روزآمدی

هم خوانی /مقایسه پذیری

دسترسی پذیری /تفسیرپذیری

تعریف: در منبع داده‌ی اداری، «شیء (object)» عنصر پایه از جامعه‌ای است که اطلاعات آن گردآوری می‌شود. شیء می‌تواند یک فرد، خانوار یا سکونت‌گاه باشد.

۱. ارزیابی خطاهای نمایندگی: به خطاهایی که ناشی از تفاوت‌های بین جامعه‌ی اداری و جامعه‌ی سرشماری است، اطلاق می‌شود.

▪ برخی نماگرهای ارزیابی تطابق اشیاء در ثبت با واحدهای هدف در سرشماری

چه تعاریف، روش‌ها و فرایندهایی برای شناسایی و درج یک شیء در منبع اداری استفاده می‌شود؟

چه قوانین یا مقرراتی برای تعریف اشیاء وجود دارد؟

آیا پوشانش اشیاء نیازهای سرشماری را برآورده می‌کند؟

▪ برخی نماگرهای ارزیابی پوشانش مجموعه‌ی اشیاء در مقایسه با جامعه‌ی هدف سرشماری

آیا شواهدی مبنی بر کم‌پوشانشی یا بیش‌پوشانشی وجود دارد که بر کارایی منبع تأثیرگذار باشد؟

۲. ارزیابی خطاهای اندازه‌گیری: به میزان عدم مطابقت اقلام داده‌ای مورد نظر نگهداری‌شده در ثبت با مفاهیم و تعاریف موضوعات سرشماری اشاره دارد. برخی از نماگرهای ارزیابی خطاهای اندازه‌گیری به شرح زیر است:

▪ آیا ثبت شامل متغیرها یا مشخصه‌های مورد نیاز برای سرشماری است؟

▪ آیا مفاهیم، تعاریف و رده‌بندی‌های اداری برای چنین متغیرهایی با آن دسته از مفاهیم، تعاریف و رده‌بندی‌های پذیرفته‌شده در سرشماری سازگاری دارد؟

▪ در صورت بروز ناسازگاری در مفاهیم، تعاریف و رده‌بندی‌های سرشماری، آیا تبدیل متغیرها یا مشخصه‌ها برای برآورده کردن الزامات سرشماری امکان‌پذیر است؟

کیفیت منبع – ارزیابی روزامدی

کیفیت منبع

مرتبط بودن

روزامدی

هم خوانی
/مقایسه پذیری

دسترسی پذیری
/تفسیرپذیری

لازم است که تفاوت بین تاریخ مرجعی که داده‌ها به آن اشاره دارند و تاریخی که داده‌ها به NSO تحویل داده می‌شوند، به حداقل برسد زیرا هر چه این تأخیر طولانی‌تر باشد، حتی در صورت درستی داده‌ها، آن‌ها کمتر مرتبط خواهند بود. برخی از نماگرهای ارزیابی روزامدی به شرح زیر است:

- تأخیر زمانی بین وقوع یک رویداد یا پدیده و ذخیره‌سازی آن در ثبت چقدر است؟
- چگونه می‌توان تاریخی که داده‌های اداری به آن اشاره دارند، با تاریخ مرجع سرشماری مقایسه کرد؟
- آیا ثبت به هنگام ارائه به NSO به‌طور کامل به‌روز شده است؟ و اگر برای موارد پیچیده‌تر تأخیر وجود داشته باشد، آیا ممکن است باعث اریبی شود؟
- فراوانی ارائه‌ی داده‌ها به NSO برای برآورده کردن نیازهای سرشماری (از قبیل به‌روزرسانی یا ایجاد افراد و سکونت‌گاه‌های جدید) چقدر است؟

کیفیت منبع – ارزیابی هم خوانی / مقایسه پذیری

کیفیت منبع

مرتبط بودن

روزآمدی

هم خوانی
/مقایسه پذیری

دسترسی پذیری
/تفسیرپذیری

- ارزیابی میزان موفقیت اتصال (یا سهولت اتصال) و ترکیب منابع اداری با سایر منابع داده‌ها به منظور استفاده در سرشماری مورد نیاز است. برخی از نماگرهای ارزیابی هم خوانی /مقایسه پذیری به شرح زیر است:
- آیا منبع شامل یک شناسانه‌ی یکتا (مانند PIN) است که با کلیدهای یکتای مورد نیاز برای اتصال سرشماری مشترک باشد؟
 - اگر چنین است، آیا شناسانه برای تمام اشیاء مربوط در منبع (مانند افراد/نشانی‌ها) موجود است یا فقط برای گروه‌های جمعیتی یا مناطق جغرافیایی خاص وجود دارد؟ در غیر این صورت
 - آیا منبع شامل ترکیب یکتایی از متغیرها (مانند نام، تاریخ تولد و نشانی) است که بتواند برای اتصال سرشماری استفاده شود؟

کیفیت منبع – ارزیابی دسترسی پذیری/تفسیر پذیری

کیفیت منبع

مرتبط بودن

روزامدی

هم خوانی
/مقایسه پذیری

دسترسی پذیری
/تفسیر پذیری

شناسایی هرگونه محدودیتی که ممکن است بر توانایی NSO در دریافت و استفاده از یک منبع اداری تأثیرگذار باشد، مهم است. برخی از نماگرهای ارزیابی دسترسی پذیری/تفسیر پذیری به شرح زیر است:

- آیا انتقال داده‌ها به سهولت انجام می‌شود؟ متولیان داده‌ها ممکن است مدل‌ها، قالب‌ها، طرح‌ها، نرم‌افزارها و سخت‌افزارهای بسیار متفاوتی را با آنچه که NSO با آن آشنایی دارد، اتخاذ کنند. این موارد شامل نحوه‌ی نگهداری و انتقال داده‌ها، از جمله اقدامات امنیتی برای انتقال است. ساختار داده‌ها نیز می‌تواند پیچیده و اندازه‌ی فایل‌ها بسیار بزرگ باشد. آگاهی NSO از چنین تفاوت‌ها و پیچیدگی‌هایی برای ارزیابی امکان دریافت و انتقال مجموعه داده‌ها به سامانه‌های NSO، امری ضروری است.

- آیا فراداده‌های واضح و جامعی وجود دارد؟ ارزیابی تفسیرپذیری به وجود و دسترسی به مستندات و فراداده‌های واضح و جامع در مورد منبع اداری مربوط می‌شود. بدون فراداده‌ها، درک و ارزیابی منبع اداری در ارتباط با کاربرد مدنظر مقدور نیست.

کیفیت داده‌های ورودی – ارزیابی هماهنگ‌سازی/اعتبارسنجی

کیفیت داده‌های
ورودی

هماهنگ‌سازی
/اعتبارسنجی

پیوندپذیری

درستی
/قابلیت اعتماد

روزآمدی

لازم است که فایل‌های داده‌های منتقل شده مطابق با فرمت خوانا در سیستم‌های NSO باشد. برخی از نماگرهای ارزیابی هماهنگ‌سازی/اعتبارسنجی به شرح زیر است:

- آیا متغیرهای مورد نظر به درستی فرمت‌دهی و مشخص شده‌اند (به‌طور مثال، اطلاعات متنی، رده‌بندی، عددی و ...)?
- آیا دوره‌ی مرجع به درستی ثبت و ارائه شده است؟
- آیا متغیرهای گردآوری شده در مرحله‌ی منبع با محتوای مورد انتظار (از پیش تعریف شده) مطابقت دارد؟

کیفیت داده‌های ورودی – ارزیابی پیوندپذیری

کیفیت داده‌های
ورودی

هماهنگ‌سازی
/اعتبارسنجی

پیوندپذیری

درستی
/قابلیت اعتماد

روزآمدی

لازم است متغیرهای هر منبع داده‌های اداری مورد استفاده در اتصال رکوردی موفق در مرحله‌ی فرایند مورد ارزیابی قرار گیرد. برخی از نماگرهای ارزیابی پیوندپذیری به شرح زیر است:

- درصد متغیرهای یکتا یا ترکیبی از متغیرهای مورد استفاده در اتصال (مانند درصد شماره‌ی شناسایی فردی (PIN) یکتا یا ترکیب سن، تاریخ تولد و نشانی)،
- خطاهای اندازه‌گیری در متغیرهای پیوندی (درصد مقادیر گمشده، مقادیر نامعتبر و ...)، و
- بروز توزیع اریب (آیا درصد زیادی از مقادیر خارج از محدوده یا نامعتبر (مانند سن ثبت‌شده‌ی ۱۲۰ سال) یا مقادیر گمشده برای متغیرهای کلید اتصال در مناطق جغرافیایی مشخص وجود دارد؟).

کیفیت داده‌های ورودی – ارزیابی درستی/قابلیت اعتماد

کیفیت داده‌های
ورودی

هماهنگ‌سازی
/اعتبارسنجی

پیوندپذیری

درستی
/قابلیت اعتماد

روزآمدی

در ارزیابی درستی داده‌های ورودی، NSO باید بین خطاهای نمایندگی (خطاهای مربوط به پوشانش جامعه‌ی هدف) و خطاهای اندازه‌گیری (خطاهای مربوط به متغیر خاص مورد نظر) تمایز قائل شود. داده‌های ورودی (در سطح کلان) با سایر منابع مانند آمارگیری ملی یا سرشماری قبلی مقایسه شود.

❖ **نماگرهای پایه برای ارزیابی خطاهای نمایندگی عبارتند از:**

- تعداد کل واحدهای دریافتی (افراد/واحدهای مسکونی) برای مقایسه با تعداد مورد انتظار، و
- درصد واحدهای تکراری.

❖ **نماگرهای پایه برای اندازه‌گیری کامل بودن متغیرهای مشخصه در مجموعه داده‌های اداری در سطح کل (مانند سن، جنس، قومیت و غیره) شامل موارد زیر است:**

- تعداد و درصد مقادیر گم‌شده در متغیرهای کلیدی (مانند تاریخ تولد و جنس)،
- تعداد و درصد مقادیر خارج از محدوده در متغیرهای کلیدی (به‌عنوان مثال، سن ثبت‌شده‌ی ۱۲۰ سال)،
- تعداد و درصد مقادیر غیر قابل قبول (به‌عنوان مثال، بر اساس جدول‌بندی تقاطعی سن و وضع زناشویی که منجر به شناسایی موارد غیر قابل قبول مانند ثبت کودک ۵ ساله به‌عنوان متأهل می‌شود)، و
- بروز فراوانی‌ها، الگوها یا نقاط دورافتاده‌ی غیر منتظره بر اساس تحلیل فراوانی/توزیعی متغیرهای کلیدی.

کیفیت داده‌های ورودی – ارزیابی روزامدی

کیفیت داده‌های
ورودی

هماهنگ‌سازی
/اعتبارسنجی

پیوندپذیری

درستی
/قابلیت اعتماد

روزامدی

با مقایسه‌ی تاریخ مرجع، تاریخ تحویل تعیین‌شده و تاریخ تحویل واقعی داده‌ها می‌توان معیارهای روزامدی را تعیین کرد. بنا بر این، دو جنبه از روزامدی را می‌توان از طریق موارد زیر ارزیابی کرد:

- تفاوت بین ثبت هرگونه تغییر در داده‌های منبع توسط متولی ثبت و تاریخ رخداد واقعی تغییر در جامعه، و
- تفاوت بین تاریخ دریافت داده‌ها توسط NSO و تاریخ دوره‌ی مرجعی که داده‌ها به آن اشاره دارند.

کیفیت فرایند – ارزیابی درستی و اتصال رکوردی

کیفیت فرایند

درستی و اتصال رکوردی

پوشانش

تصحیح مغایرت

ویرایش و جانمایی

متداولترین نماگرها برای ارزیابی کیفیت اتصال عبارتند از:

- نسبت رکوردهای پیوندی (رکوردهایی که قابلیت پیوند دارند)، و غیر پیوندی (رکوردهایی که قابلیت پیوند ندارند)،
- تعداد و درصد کلیدهای اتصال تکراری،
- عدم وجود کلیدهای اتصال یا تعداد مقادیر گمشده یا غیر قابل قبول،
- مقایسه‌ی توزیع ویژگی‌های رکوردهای پیوندی و غیر پیوندی برای مثال به تفکیک متغیرهایی مانند سن، جنس و قومیت. تفاوت در ویژگی‌ها می‌تواند به اربابی ناشی از خطای اتصال اشاره داشته باشد.

کیفیت فرایند – ارزیابی کیفیت ثبت آماری جمعیت (SPR)

❖ ارزیابی پوشانش با استفاده از نشانه‌های زندگی ((Signs of Life (SOL))

روش نشانه‌های زندگی به‌عنوان ابزاری برای به حداقل رساندن بیش‌پوشان‌شی مورد استفاده قرار می‌گیرد. در این روش تعدادی از قاعده‌های فعالیت اتخاذ می‌شود تا اطمینان حاصل شود که تنها افرادی که زنده هستند و مجموعه‌ای از معیارهای اقامت از پیش تعریف‌شده را برآورده می‌کنند، در سرشماری گنجانده شوند. این روش معمولاً شامل ایجاد برخی مفروضات است که مشخص می‌کند چه کسی فعال و چه کسی غیر فعال است. انتخاب نشانه‌های نماگرهای فعالیت (قاعده‌های تصمیم‌گیری) باید با ارزیابی کیفیت در مرحله‌ی منبع و داده‌های ورودی، از جمله مشاوره با متولیان داده‌ها، اعتبارسنجی متقابل بین منابع و در طول زمان و نظر کارشناسی انجام شود (مثال، کشور اسپانیا و لهستان).

برای مثال، در اسپانیا نشانه‌های زندگی با استفاده از داده‌های چندین منابع اداری از جمله ثبت مالیات، پرونده‌های تأمین اجتماعی و پایگاه‌داده‌ی ملی بیکاری و جابه‌جایی‌های شناسایی‌شده در ثبت جمعیت محلی تعیین می‌شوند. این منابع اداری در سطح فردی و خانواری به هم مرتبط هستند و افرادی که سطح آستانه‌ی نشانه‌های زندگی خاص که نشان دهنده‌ی حضور فعال آن‌ها در کشور است را به‌دست آوردند، در شمارش جمعیت گنجانده می‌شوند و بقیه‌ی افراد که غیر فعال شناخته شده‌اند، در شمارش در نظر گرفته نمی‌شوند.

❖ ارزیابی پوشانش با استفاده از آمارگیری‌های مستقل

علاوه بر رویکرد نشانه‌های زندگی، روش‌های دیگری نیز برای برآورد پوشانش (و همچنین خطای محتوای) سرشماری‌ها وجود دارد. این روش‌ها عبارتند از تحلیل‌های جمعیت‌شناختی، مقایسه با داده‌های سایر منابع از جمله سرشماری‌های پیشین، آمارگیری‌های خانواری جاری و آمارگیری‌های پس‌سرشماری.

کیفیت فرایند

درستی و اتصال
رکوردی

پوشانش

تصحیح مغایرت

ویرایش و جان‌هی

کیفیت فرایند – تصحیح مغایرت

کیفیت فرایند

درستی و اتصال
رکوردی

پوشانش

تصحیح مغایرت

ویرایش و جانمایی

- هنگام استفاده از داده‌های منابع مختلف، روش‌هایی برای ارزیابی کیفیت متغیرهایی که به متغیر یکسانی در منابع مختلف اشاره دارد، مورد نیاز است. از آنجا که یک فرد ممکن است در منابع داده‌های مختلف با اطلاعات مغایر ظاهر شود، قاعده‌هایی (مانند روش SOL که پیش از این توضیح داده شد) لازم است تا مشخص شود کدام مقدار درست است.
- ممکن است نیاز به رفع مشخصات متناقض نشانی‌های ثبت شده در ثبت‌های مختلف و شناسایی درست‌ترین اطلاعات وجود داشته باشد. چنین تفاوت‌هایی ممکن است به‌عنوان مثال از تأخیر افراد در گزارش‌دهی تغییر نشانی یا گزارش‌دهی نشانی منزل دوم ناشی شود. اطلاعات نشانی متناقض (یا متعدد) و هر تصمیم مربوط ممکن است منجر به کم‌پوشانشی در برخی مناطق و بیش‌پوشانشی در برخی مناطق دیگر شود. بنا بر این NSO باید تصمیم بگیرد که اطلاعات سرشماری به کدام نشانی منتسب شود (به‌طور مثال نشانی که اخیراً گزارش شده باشد، می‌تواند در نظر گرفته شود).

کیفیت فرایند - ارزیابی فرایند ویرایش و جانهی

کیفیت فرایند

درستی و اتصال
رکوردی

پوشانش

تصحیح مغایرت

ویرایش و جانهی

ویرایش و جانهی یک فرایند تکراری برای بازبینی داده‌ها به منظور تصحیح هر گونه خطای ناشی از واردسازی داده‌های نامعتبر، ناسازگار یا غیر قابل قبول و جایگزینی این موارد و داده‌های گمشده با مقادیر قابل قبول است. برای پایش روش‌های جانهی و کاربرد قاعده‌های ویرایش می‌توان از نماگرهای زیر استفاده کرد:

۱. نرخ شکست ویرایش (edit failure rate): این نرخ از تقسیم تعداد خطاهای برطرف‌شده بر تعداد کل خطاها (خطاهای برطرف‌شده و برطرف‌نشده) در فایل داده‌ها به دست می‌آید. این نماگر بین مقادیر ۰ و ۱ متغیر است. مقدار بالا در این نماگر به این معنی است که قاعده‌ی ویرایش به درستی به عنوان معیار خطایابی تعریف شده است و مقدار کم این نماگر به این معنی است که قاعده‌ی ویرایش با شکست مواجه شده و باید بهبود یابد،

۲. نرخ جانهی: این نرخ نشان‌دهنده‌ی درصد رکوردهای جانهی‌شده در کل رکوردها است. این نماگر باید برای هر متغیر محاسبه شود. نرخ‌های جانهی، نماگرهای کلیدی کیفیت برای سرشماری هستند،

۳. شاخص عدم تشابه (dissimilarity index): این شاخص نشان‌دهنده‌ی میزان تغییر توزیع درصد متغیر قبل از جانهی و توزیع درصد متغیر بعد از جانهی است. محدوده‌ی این شاخص از ۰ تا ۱۰۰ متغیر است که در آن ۱۰۰ به معنای حداکثر عدم تشابه است.

کیفیت برون داد- ارزیابی مرتبط بودن

کیفیت برون داد

مرتبط بودن

درستی

روزامدی

هم خوانی /مقایسه پذیری

دسترسی پذیری /تفسیر پذیری

برای آن دسته از NSOهایی که به روش شناسی گردآوری داده های سرشماری ترکیبی یا ثبتی مبنا روی آورده اند، ارزیابی کیفیت برون داد برای تعیین این که آیا این گذار بر کیفیت کلی برون داده ها تأثیر گذاشته است یا خیر، بسیار مهم است. ارزیابی کیفیت برون داد نباید محدود به برآورد عدم حتمیت کل (بعد درستی) شود، بلکه باید شامل ارزیابی در تمام ابعاد دیگر کیفیت نیز باشد.

سه رویکرد رایج برای ارزیابی مرتبط بودن برون داد عبارتند از:

- پرسش مستقیم از کاربران از طریق اجرای آمارگیری نیازهای کاربران (قبل از سرشماری) یا آمارگیری رضایت مندی کاربران (پس از سرشماری)،
- ایجاد سازوکارهای بازخورد کاربران در فرایند سرشماری، و
- تحلیل کاربرد داده های سرشماری (به عنوان مثال از طریق بررسی گزارش ها و مقالات منتشر شده).

کیفیت برون داد – ارزیابی درستی

کیفیت برون داد

مرتبط بودن

درستی

روزامدی

هم خوانی
/مقایسه پذیری

دسترسی پذیری
/تفسیر پذیری

- اندازه گیری درستی کل شمارش نهایی سرشماری به طور سنتی بر اساس برآورد نظام دوگان (Dual System Estimation) است که معمولاً شامل اجرای یک آمارگیری پس سرشماری و سپس اتخاذ روش های گیر و بازگیر (Capture-Recapture) برای برآورد کم پوشانشی و بیش پوشانشی است.

- برخی از NSO ها به طور سنتی آمارگیری های نمونه ای با حجم کمتر برای واریسی کیفیت پس سرشماری را انجام می دهند که در آن داده ها در مورد تمام سؤالات سرشماری گردآوری می شوند و سپس با مجموعه داده های آماری جور می شوند تا خطای پاسخ گو اندازه گیری شود.

- در روش شناسی سرشماری ثبتی مبنا، روش های سنتی برای تعیین کیفیت و پوشانش کل تا حدودی هنوز قابل اجرا هستند. با این وجود، در مورد شمارش های سرشماری که صرفاً از ثبت های اداری تهیه می شوند، روش های جدید یا بازبینی شده ضروری است.

- برخی از NSO ها ممکن است راهبردی را اتخاذ کنند که نیازمند اتکای کمتری به رویکردهای سنتی (مانند برآورد نظام دوگان DSE) برای ارزیابی پوشانش است و بیشتر بر کیفیت نگهداری و بهنگام سازی ثبت های جمعیت متکی است. به عنوان مثال، بسیاری از کشورها تلاش کرده اند تا روش هایی را توسعه دهند که هدف آن ها کاهش بیش پوشانشی از طریق روش های «نشانه های زندگی» است که در بخش قبل ذکر شده است. گرچه چنین روش های جدید و بازبینی شده با پیشرفت روزافزونی در میان NSO ها مورد توجه قرار گرفته است، اما استفاده از روش های سنتی DSE نیز می تواند سودمند باشد.

کیفیت برون داد- ارزیابی روزامدی

کیفیت برون داد

مرتبط بودن

درستی

روزامدی

هم خوانی
/مقایسه پذیری

دسترسی پذیری
/تفسیر پذیری

- روزامدی به فاصله‌ی زمانی بین تاریخ داده‌های سرشماری و تاریخ انتشار داده‌ها اشاره دارد.
- روزامدی برون داده‌ها تا حد زیادی با روزامدی تحویل داده‌های ورودی و زمان صرف شده توسط NSO برای انجام پردازش کامل داده‌ها تعیین می‌شود. اغلب بین روزامدی و درستی رابطه‌ای وجود دارد و به‌طور کلی هر چه NSO زمان بیشتری را برای اطمینان از درستی داده‌ها در طول پردازش اختصاص دهد، تحویل برون داده‌ها به کاربر بیشتر طول می‌کشد.
- شاخص‌های کمی را می‌توان برای اندازه‌گیری روزامدی نتایج نهایی شامل هرگونه تأخیر بین گردآوری داده‌ها، دریافت داده‌ها، اتصال داده‌ها و انتشار آمارها به‌کار برد. به‌عنوان مثال، روزامدی کل می‌تواند به‌صورت دوره‌ی زمانی از پایان دوره‌ی مرجع تا دریافت داده‌های اداری، تقسیم بر زمان از پایان دوره‌ی مرجع تا تاریخ انتشار، ضرب در ۱۰۰ درصد محاسبه شود.
- بعد «به‌موقع بودن» را می‌توان به‌سادگی با تعیین تفاوت تاریخ از پیش تعیین شده‌ی تحویل برون داده‌ها و تاریخ واقعی تحویل ارزیابی کرد.

کیفیت برون داد- ارزیابی هم خوانی /مقایسه پذیری

کیفیت برون داد

مرتبط بودن

درستی

روزامدی

هم خوانی
/مقایسه پذیری

دسترسی پذیری
/تفسیر پذیری

- برآوردهای تولیدشده توسط داده‌های اداری باید با مشخصه‌های شناخته‌شده‌ی جمعیت هم از لحاظ طولی (از طریق مقایسه‌ی یک سرشماری با سرشماری دیگر) و هم از لحاظ جغرافیایی (از طریق مقایسه‌ی یک ناحیه با ناحیه‌ی دیگر) با برون داده‌های حاصل از سایر عملیات داده‌های آماری NSO سازگار باشد.
- تحلیل‌های جمعیت‌شناختی یک رویکرد خاص است که می‌تواند برای ارزیابی هم خوانی و مقایسه‌پذیری (و همچنین درستی) برون داده‌های سرشماری به‌کار رود. تحلیل‌های جمعیت‌شناختی شامل مقایسه‌های سیستماتیک، تعیین آستانه‌های مقبولیت و درک هر گونه تفاوت معنی‌دار است.

کیفیت برون داد- ارزیابی دسترسی پذیری / تفسیر پذیری

کیفیت برون داد

مرتبط بودن

درستی

روزآمدی

هم خوانی
/ مقایسه پذیری

دسترسی پذیری
/ تفسیر پذیری

- دسترسی پذیری هم به سهولت فهم و تشخیص ماهیت داده ها توسط کاربر و هم به مناسب بودن شیوه ها و قالب هایی که داده ها را می توان در دسترس کاربر قرار داد، اشاره دارد.
- تفسیر پذیری یا وضوح به در دسترس بودن هر گونه اطلاعات تکمیلی یا فراداده های ضروری برای کمک به کاربر در تفسیر و درک برون داده های آماری اشاره دارد.
- اجرای آمارگیری رضایت مندی کاربر توسط NSO و یا تشکیل کارگروه بازبینی کارشناسی پس از تکمیل برنامه ی برون داد برای ارزیابی میزان دسترسی پذیری و وضوح (و همچنین مرتبط بودن) از دیدگاه کاربران مفید خواهد بود.

بخش سوم: مروری بر مطالعات تطبیقی



استرالیا – استفاده از داده‌های اداری برای پشتیبانی از عملیات سرشماری

استفاده از داده‌های اداری برای سرشماری ۲۰۱۶ موجب تسهیل در موارد زیر شد:

۱. به کارگیری داده‌های اداری مربوط به ساختمان‌ها یا نشانی‌ها همراه با سایر روش‌های گردآوری داده‌ها و فرایندهای راستی آزمایی موجب ایجاد ثبت ملی نشانی برای کاربرد در سرشماری‌های بعدی شد. ثبت نشانی با استفاده از فایل ملی نشانی کدبندی جغرافیایی به‌عنوان پایه‌ی آن شکل گرفت و سپس علاوه بر تحلیل دیگر داده‌های موجود، با استفاده از اطلاعات به‌دست آمده از مراجعه به هر نشانی ساخته شد. این ثبت پایه، نشانی‌هایی را تشکیل می‌داد که اطلاعات به آن‌ها پست می‌شد.
۲. نامه‌ی دستورالعمل که نحوه‌ی تکمیل سرشماری برخط یا نحوه‌ی درخواست پرسش‌نامه‌ی کاغذی را شرح می‌داد، از طریق پست به این نشانی‌ها ارسال شد. سپس پرسش‌نامه‌های کاغذی ارسال و از طریق پست استرالیا بازگردانده شدند و فقط زمانی که استفاده از خدمات پستی میسر نبود یا پرسش‌نامه‌های سرشماری بازگردانده نشده بودند، مأموران میدانی سرشماری به آن خانه‌ها مراجعه کردند.
۳. بنا بر این استفاده از داده‌های اداری منجر به افزایش استفاده از پرسش‌نامه‌های سرشماری برخط به عنوان روش پیش فرض برای اکثر خانوارهای استرالیایی شد.

توسعه‌ی استفاده از داده‌های اداری برای سرشماری ۲۰۲۱ استرالیا

- استفاده از داده‌های اداری به منظور کمک‌رسانی به جوامع و گروه‌های خاص در حین سرشماری: دفتر آمار استرالیا، اغلب از ترکیب داده‌های اداری با داده‌های سرشماری قبلی، برای شناسایی و درک بهتر جوامع یا گروه‌هایی استفاده می‌کند که در گذشته نرخ پاسخ‌گویی کمتر از متوسط داشته‌اند. به عنوان مثال، با شناسایی مناطقی که سالمندان زیادی داشتند که ممکن بود در پر کردن پرسش‌نامه‌های برخط با مشکل مواجه شده باشند، دفتر آمار استرالیا می‌توانست کارکنانی را به منظور کمک‌رسانی به گروه هدف در تکمیل پرسش‌نامه‌ی سرشماری اعزام کند.
- استفاده از داده‌های اداری برای کمک‌رسانی به بهبود شمارش در سرشماری از طریق موارد زیر:
 - ✓ سرشماری استرالیا یک سرشماری حاضرشمار (De Jure) است. بنا بر این یک سؤال کلیدی در سرشماری استرالیا این است که آیا خانه یا سکونت‌گاه بدون پاسخ، خانه‌ای است که در شب قبل، خالی از سکنه بوده است یا این که شخصی در خانه بوده اما پاسخی از او دریافت نشده است.
 - ✓ ابتدا از داده‌های اداری برای ارزیابی خانه‌های خالی در شب سرشماری استفاده شد. داده‌های مورد استفاده شامل داده‌های خدمات دولتی از بخش‌هایی مانند حوزه‌ی سلامت و بهداشت، خدمات اجتماعی و مالیات بود. از داده‌های اداری بخش‌های خصوصی مانند داده‌های مصرف برق از طریق توزیع‌کنندگان برق سطح ناحیه (و نه در سطح فردی) استفاده شد.
 - ✓ هنگامی که مشخص شد خانه‌ای تصرف شده و هیچ داده‌ای در دسترس نیست، از داده‌های اداری برای انتخاب بهترین روش برای تعیین تعداد افراد ساکن در آن خانه استفاده شد. این کار، با کپی‌برداری داده‌های اداری از خانه‌ای که دارای ویژگی‌های مشترک با خانه‌ی موردنظر بود، انجام شد.
- آمادگی به هنگام پیشامدهای غیر منتظره (مانند آتش‌سوزی مراتع) و استفاده از داده‌های اداری به منظور پر کردن شکاف موجود در داده‌های سرشماری
 ۱. زمانی که اطلاعات سرشماری برای خانوارهای درگیر پیشامد غیر منتظره گمشده باشد، امکان استفاده از داده‌های ادارای فراهم شد.
 ۲. اطلاعات دقیق‌تر سرشماری برای خانوارهای درگیر پیشامد غیر منتظره در مورد متغیرهایی مانند تاریخ تولد و جنس که در طول زمان تغییر نمی‌کنند، توسط منابع اداری یا داده‌های سرشماری قبلی می‌توانست تکمیل شود.

اندونزی – استفاده از داده‌های اداری برای خانوارهای از پیش فهرست‌شده

▪ سرشماری ۲۰۲۰ – معرفی داده‌های اداری و فناوری‌های جدید در مراحل مختلف اجرای سرشماری از قبیل:

۱. استفاده از داده‌های ثبت احوال به‌عنوان ورودی در فاز پیش‌فهرست‌برداری: گرچه داده‌های جمعیتی حاصل از نظام ثبت احوال برای استفاده در تولید متغیرهای سرشماری از کیفیت کافی برخوردار نبودند، اما برای عملیات سرشماری بسیار مفید بودند. کاربرد عمده‌ی داده‌های اداری در فاز پیش‌فهرست‌برداری بود. هنگامی‌که داده‌های جمعیتی حاصل از ثبت احوال توسط اداره‌ی آمار اندونزی دریافت شد، این فایل با فایل جمعیتی موجود در اداره‌ی آمار اندونزی که بر اساس سرشماری قبلی بود، مطابقت داده شد. حتی اگر همه‌ی شهروندان این کشور، شناسه‌ی یکتا داشته باشند، کیفیت پایین این متغیر به این معنی است که اداره‌ی آمار اندونزی مجبور به کاربست رویکردهای جورسازی‌های مختلف است. اولین گام در رویکرد جورسازی، استفاده از رویکرد قطعی بود که در آن شناسه‌ی یکتا و نام افراد در دو منبع یکسان بود و انجام جورسازی را عملی می‌کرد. رویکرد دوم استفاده از رویکرد احتمالاتی بود که در آن متغیرهای بیشتری برای جورسازی استفاده می‌شد. علاوه بر این، جورسازی برای مناطق اداری به‌منظور تسهیل تهیه‌ی فهرست جمعیت متناسب با مکان‌ها انجام شد.

۲. مرحله‌ی اول گردآوری اطلاعات - سرشماری برخط: اطلاعات گردآوری شده در فاز پیش‌فهرست‌برداری منجر به ثبت موقت جمعیت آماری با اطلاعات پایه در سطح فردی شد. در مرحله‌ی دوم سرشماری برخط اجرا شد. در این مرحله، افراد تشویق شدند تا اطلاعات سرشماری خود را از طریق پرسش‌نامه‌ی برخط تکمیل کنند. این پرسش‌نامه‌ی برخط با اطلاعات ثبت آماری جمعیت از قبل تکمیل شده بود و از پاسخ‌گویان درخواست شد تا اطلاعات را تأیید یا به‌روز کنند (یا در صورت گمشدگی اطلاعات را اضافه کنند). بر اساس سرشماری برخط، ثبت آماری جمعیت به‌روز شده تولید شد.

۳. مرحله‌ی دوم گردآوری داده‌ها - تأیید در سطح اداری محلی: در این مرحله، ثبت آماری جمعیت به‌روز شده در سطح محلی توسط کارکنان مرکز آمار اندونزی با همکاری مقامات محلی تأیید شد. از مقامات محلی برای تأیید فهرست خانوارهای منطقه‌ی اداری آن‌ها درخواست همکاری شد. خانوارهایی که مقامات محلی از وجود آن‌ها مطمئن نبودند، علامت‌گذاری شدند. سپس این خانوارها در عملیات میدان شناسایی شدند و آمارگیران میدانی به آن‌ها مراجعه کردند. پس از تأیید حضور اعضای خانوار، یک پرسش‌نامه‌ی سرشماری کاغذی تحویل داده شد تا بعد از تکمیل از خانوارها دریافت شود.

۴. مرحله‌ی سوم - شمارش میدانی: باقی خانوارهای ثبت آماری جمعیت تأیید شده که به پرسش‌نامه‌ی برخط پاسخ نداده بودند یا در مرحله‌ی دوم مشکوک شناسایی شده بودند و برای تحویل پرسش‌نامه به آن‌ها مراجعه شده بود، در مرحله‌ی نهایی گردآوری اطلاعات از طریق فرم و پرسش‌نامه‌ی کوتاه به آن‌ها مراجعه شد. سپس این خانوارها با استفاده از مصاحبه‌ی شخصی یا رانه‌یار (CAPI) یا مصاحبه‌ی قلم و کاغذی (PAPI) توسط آمارگیران میدانی مصاحبه شدند. پس از گردآوری داده‌ها از طریق تکمیل فرم و پرسش‌نامه‌ی کوتاه، آمارگیری پس‌سرشماری نیز انجام شد.

۵. فرم‌ها و پرسش‌نامه‌های طولانی - آمارگیری نمونه‌ای: در حالی که شمارش از طریق فرم و پرسش‌نامه‌ی کوتاه مستلزم کسب اطلاعات از کل جمعیت و تولید داده‌های پایه‌ی جمعیت بود، اما آمارگیری نمونه‌ای از طریق فرم و پرسش‌نامه‌ی طولانی که اطلاعات دقیق‌تری مانند باروری، مرگ و میر و مهاجرت ارائه می‌دهد، اجرا شد. حدود ۵ درصد از جمعیت، یا حدود ۳ میلیون خانوار، در این آمارگیری شرکت کردند. این آمارگیری که در سال ۲۰۲۱ انجام شد، برآوردهای نمایان‌گر را در سطح زیر منطقه‌ای ارائه کرد. برای گردآوری داده‌ها از هر دو روش CAPI و PAPI استفاده شد.

جمهوری کره – اقدامات انجام شده در حرکت به سمت سرشماری ترکیبی ۲۰۲۰

- جمهوری کره اولین سرشماری ترکیبی را در سال ۲۰۱۵ اجرا کرد.
- استفاده از ۲۵ ثبت اداری از ۱۴ مؤسسه دولتی: برای انجام این کار، بسیاری از ثبت‌های مختلف باید در سطح واحد متصل می‌شدند. شماره‌های شناسایی فردی در ثبت‌ها به اعداد (۱۰ رقمی) تبدیل شدند که به عنوان کلید جورسازی عمل می‌کردند. این کلید جورسازی فقط برای اهداف آماری ایجاد شد.
- شناسایی ناسازگاری‌های داده‌های اداری و سرشماری نفوس موجب بهبود سرشماری ترکیبی ۲۰۲۰ شد:
 ۱. ثبت جمعیت (یا پایگاه داده‌ی ثبت مقیمان (RRR)) شامل افرادی است که ممکن است برای کار، تحصیل یا به دلایل دیگر به خارج از کشور نقل مکان کرده باشند. این افراد باید از داده‌های مربوط به مهاجرت حذف می‌شدند تا افراد حاضر در جمهوری کره بهتر منعکس شوند. علاوه بر این، تأخیرهایی در گزارش‌دهی ولادت و وفات وجود داشت که موجب بروز اختلافاتی بین داده‌های اداری و سرشماری شده بود.
 ۲. مقایسه‌ی دیگری برای تعداد خانوارها انجام شد. داده‌های اداری، ۲/۵ میلیون خانوار بیشتر از داده‌های سرشماری نفوس بود. این تفاوت به این دلیل بود که برخی از اعضای این خانوارها، به دلایل تحصیل در مدرسه و داشتن مسکن، در نشانی‌هایی متفاوت از آنچه که در ثبت‌های اداری گزارش شده بود، سکونت داشتند.
 ۳. بررسی مطابقت بین محل اقامت واقعی (موارد شمارش شده در سرشماری) و محل اقامت ثبت شده (در ثبت‌های اداری) انجام شد.
 ۴. از داده‌های بیمه‌ی بازنشستگی و اشتغال برای تصحیح مواردی که تغییر محل کار افراد مشخص نشده بود، استفاده شد و از قلم‌افتادگی‌ها، تأخیرها و خطاهای احتمالی در داده‌های اداری با استفاده از داده‌های پیشین مشخص شد.
- استفاده از روش‌های مختلف جورسازی بین منابع داده‌های اداری و بین داده‌های اداری و سرشماری

چالش‌های کلیدی در حرکت به سمت سرشماری ثبتی‌مبنای به‌موقع و معتبر در جمهوری کره

۱. نهادینه‌سازی دسترسی به داده‌های اداری ارائه‌شده توسط متولیان آن‌ها برای مرکز آمار کره،
۲. بازبینی کیفیت داده‌های اداری،
۳. استانداردسازی داده‌های اداری در منابع مختلف،
۴. بهبود روزآمدی داده‌های اداری، و
۵. حفاظت مستمر اطلاعات فردی.

ترکیه – اقدامات طرح‌ریزی شده به سمت سرشماری ثبتی مبنا

مراحل ایجاد نظام ثبت جمعیت نشانی مبنا (ABPRS) (Address-Based Population Registration System):

- ایجاد نظام ثبت احوال متمرکز به نام مرنیس در اوایل دهه‌ی ۲۰۰۰ و اختصاص شماره‌ی شناسایی یکتای ۱۱ رقمی به تمام شهروندان، از جمله افرادی که در قید حیات نبودند.
- در ابتدای کار، مؤسسه‌ی آمار ترکیه پایگاه‌داده‌ی نشانی ملی NAD (National Address Database) را ایجاد کرد که شامل تمام نشانی‌های کشور با برخی اطلاعات پایه برای توصیف هر نشانی یا سکونت‌گاه بود.
- سپس مؤسسه‌ی آمار ترکیه، شمارش میدانی کامل را برای تعیین محل اقامت معمول هر فرد به همراه شماره‌ی شناسایی حاصل از پایگاه‌داده‌ی مرنیس انجام داد. نتایج حاصل از این شمارش با NAD منجر به ساخت ABPRS شد.

ویژگی‌های پایگاه‌داده‌ی ABPRS:

۱. این پایگاه‌داده نه تنها یک پایگاه‌داده‌ی آماری است، بلکه منبع داده‌ی اداری است که توسط سایر سازمان‌های دولتی نیز استفاده می‌شود.
۲. این پایگاه‌داده به‌روز است زیرا طبق قانون، هرگونه تغییر نشانی باید ظرف مدت ۲۰ روز به ثبت اطلاع داده شود.

سرشماری ۲۰۱۱ – معرفی رویکرد ترکیبی در ترکیه

- سرشماری ۲۰۱۱ اولین فرصت برای استفاده از نظام ثبت جمعیت نشانی مبنا (ABPRS) برای عملیات سرشماری بود. این پایگاه داده توانست تقریباً ۲۰ درصد از متغیرهای سرشماری را به طور مستقیم از ثبت تولید کند، ۸۰ درصد باقی مانده باید با استفاده از یک آمارگیری نمونه‌ای مقیاس بزرگ گردآوری می‌شد.
- داده‌های ثبتی حاصل از پایگاه داده‌ی نشانی ملی (NAD) نیز برای فهرست‌برداری خانوارها و حوزه‌های سرشماری برای عملیات میدانی استفاده شد.
- همچنین، در سال ۲۰۱۱ برای اولین بار از تبلت‌ها برای سرشماری استفاده شد، به طوری که حدود ۶۰ درصد مصاحبه‌ها با استفاده از CAPI و ۴۰ درصد باقی مانده با استفاده از PAPI انجام شد.
- نتیجه‌ی استفاده‌ی ترکیبی از داده‌های ثبتی و کاربرد تبلت‌ها در حین گردآوری داده‌های میدانی، کاهش فاصله‌ی زمانی بین گردآوری داده‌ها و انتشار نتایج سرشماری بود.

گسترش استفاده از داده‌های ثبتی در سرشماری ۲۰۲۱ ترکیه

بین سرشماری ۲۰۱۱ و آغاز طرح‌ریزی برای سرشماری ۲۰۲۱، چندین رویداد مهم رخ داد:

- در سال ۲۰۱۲، پایگاه‌داده‌ی نشانی اقامت‌گاه‌های مؤسسه‌ای (Institutional Living Quarters Address Database) ایجاد شد که گردآوری داده‌ها در مورد گروه‌های جمعیتی ساکن در این مؤسسات را آسان‌تر کرد. چندین ثبت دیگر نیز ایجاد شدند یا توسعه یافتند.
- همچنین، پایگاه‌داده‌ی نشانی ملی (NAD) با افزودن اطلاعات مختصات به نشانی‌ها به نظام ثبت نشانی مکانی تبدیل شد.
- به‌منظور تسهیل استفاده از منابع اداری دیگر برای اهداف آماری، مؤسسه‌ی آمار ترکیه پایگاه‌داده‌ی مشخصه‌های جمعیت را توسعه داد.
- با ترکیب تمام ثبت‌های تکمیلی در پایگاه‌داده‌ی مشخصه‌های جمعیت با نظام ثبت جمعیت نشانی‌مبنا (ABPBS)، مؤسسه‌ی آمار ترکیه حدود ۷۰ درصد از متغیرهای سرشماری سال ۲۰۲۱ را با استفاده از داده‌های ثبتی تولید کرد. برای ۳۰ درصد باقی‌مانده از متغیرهای سرشماری در مورد مشخصه‌های ساختمان‌ها و سکونت‌گاه‌ها، مؤسسه‌ی آمار ترکیه یک آمارگیری موردی بین اکتبر ۲۰۲۱ و مارس ۲۰۲۲ انجام داد.
- مؤسسه‌ی آمار ترکیه از فن CATI برای این آمارگیری استفاده کرد. مدت زمان برآوردشده‌ی آمارگیری برای هر خانوار شش دقیقه بود که نشان‌دهنده‌ی کاهش بار پاسخ‌گویی به‌دلیل گردآوری بیشتر اطلاعات به‌طور مستقیم از ثبت‌ها است.

چالش‌های کلیدی برای استفاده از داده‌های اداری در تولید آمار (از جمله سرشماری‌ها) در ترکیه

- فقدان فراداده‌ها،
- رده‌بندی‌های متفاوت توسط متولیان داده‌های مختلف،
- مقادیر گم‌شده در ثبت‌ها،
- تفاوت در تعاریف،
- اختلاف در تاریخ‌های مرجع بین منابع داده‌ها،
- وجود رکوردهای چندگانه (موارد تکراری)، و
- فقدان زیرساخت کافی فناوری اطلاعات.



- ✓ داده‌های اداری می‌توانند به روش‌های مختلفی برای بهبود کیفیت و کارایی سرشماری ترکیبی مورد استفاده قرار گیرند.
- ✓ تصمیم‌گیری درباره‌ی استفاده از یک ثبت اداری مشخص برای تولید داده‌های سرشماری بایستی بر اساس ارزیابی کیفیت منبع داده‌های اداری و کیفیت داده‌های ورودی صورت پذیرد.
- ✓ تمام مراحل ارزیابی کیفیت (منبع، داده‌های ورودی، فرایند و برون‌داد) بایستی به‌خوبی مستندسازی و برای آگاهی کاربران منتشر شود.
- ✓ دو ویژگی کلیدی در فرایند مدرن‌سازی در استرالیا، استفاده از داده‌های اداری برای بهبود فرایندهای مختلف سرشماری و استفاده از گردآوری داده‌های برخط برای کاهش نیاز به بازدیدهای میدانی است. سرشماری استرالیا می‌تواند مثال مفیدی برای کشورهایی باشد که قصد ندارند ثبت جمعیت ایجاد کنند یا اطلاعات سرشماری را با استفاده از داده‌های اداری گردآوری کنند، اما مایلند کارایی سرشماری سنتی خود را با استفاده از داده‌های اداری موجود بهبود بخشند.
- ✓ اندونزی توانسته است از مزایای داده‌های اداری جمعیت برای کاهش هزینه‌ها و بهبود اثربخشی عملیات سرشماری استفاده کند. گرچه داده‌های اداری از کیفیت کافی برای ارائه‌ی مستقیم متغیرهای سرشماری در اندونزی برخوردار نبودند، اما از آن‌ها برای بهبود و سرعت بخشیدن به فهرست‌برداری خانوارها استفاده شد و با ترکیب داده‌های اداری و سرشماری برخط، به‌طور قابل توجهی نیاز به سرشماری میدانی کاهش یافت.
- ✓ با حرکت به سمت سرشماری ترکیبی، اکنون مرکز آمار جمهوری کره قادر است ۱۳ قلم سرشماری را به‌صورت سالانه منتشر کند. همچنین، مرکز آمار جمهوری کره به دنبال گسترش تعداد اقلام سرشماری حاصل از داده‌های ثبتی، در آمار نفوس و مسکن است. رویکرد مورد استفاده در جمهوری کره، مثال خوبی برای کشورهایی است که منابع داده‌های اداری تثبیت‌شده دارند و به دنبال حرکت به سمت سرشماری ترکیبی هستند.
- ✓ ترکیه نیز نمونه‌ای از کشورهایی است که از طریق رویکرد ترکیبی در حال انتقال طرح‌ریزی‌شده از سرشماری سنتی به سرشماری ثبتی‌مبنای کامل است. ترکیه چارچوب قانونی، ظرفیت سازمانی و زیرساخت‌های فناوری اطلاعات و سایر زیرساخت‌های لازم را به‌طور نظام‌مند برای استفاده از منابع داده‌های اداری موجود در کشور ایجاد کرده‌اند. علاوه بر این، ارتباط نزدیکی با متولیان منابع داده‌های اداری مختلف برقرار کرده‌اند تا هم به این منابع دسترسی داشته باشند و هم در صورت لزوم، برای بهبود کیفیت ثبت‌ها به متولیان داده‌ها یاری رسانند.

مرجع‌های اصلی

- United Nations Statistics Division (UNSD). (2022). *Handbook on Registers-Based Population and Housing Censuses*. <https://unstats.un.org/unsd/statcom/53rd-session/documents/BG-3e-Handbook-E.pdf>.
- United Nations Economic and Social Commission for Asia and the Pacific (ESCAP). (2022). *Emerging Trends in Census Approaches in Asia and the Pacific: with Country Examples*.
- ترابی کهلان، پریا و کلهری ندرآبادی، لیدا. (۱۴۰۳). توصیه‌های بین‌المللی در به‌کارگیری داده‌های ثبتی و اداری در سرشماری نفوس و مسکن، پژوهشکده‌ی آمار.
- ترابی کهلان، پریا و کلهری ندرآبادی، لیدا. (۱۴۰۳). کاربرد داده‌های ثبتی و اداری در سرشماری نفوس و مسکن، مطالعه‌ی موردی: برخی کشورهای منطقه‌ی آسیا و اقیانوسیه، مجله‌ی اندیشه‌ی آماری، ۵۶ (۲)، ۵۵-۶۸.
- ترابی کهلان، پریا و کلهری ندرآبادی، لیدا. (۱۴۰۲). معرفی ابزارها و نماگرهای کلیدی ارزیابی کیفیت ثبت‌ها و داده‌های اداری برای استفاده در سرشماری نفوس و مسکن، مجله‌ی اندیشه‌ی آماری، ۵۵ (۱)، ۴۱-۲۷.



www.srtc.ac.ir

با سپاس از توجه شما

