



به نام او که به شمارش همه موجودات آگاه است

دستورالعمل ارزیابی کیفیت ثبت اداری و قابلیت تولید ثبت آماری

ویرایش سوم

۲۱ شهریور ۱۴۰۰

کمیته طراحی و روش‌شناسی

فهرست مطالب

پیشگفتار.....	۲
۱- مقدمه.....	۳
۲- ارزیابی ثبت اداری و بررسی قابلیت استفاده در تولید ثبت آماری ($q_{\alpha 1}$, $q_{\alpha 2}$).....	۳
۲-۱- کیفیت ثبت داده‌ها (q).....	۴
۲-۱-۱- شاخص‌های ارزیابی «کیفیت ثبت داده‌ها».....	۴
۲-۱-۲- راهنمای ارزیابی «کیفیت ثبت داده‌ها».....	۶
۲-۲- کیفیت فایل داده‌های دریافتی (q_1).....	۱۰
۲-۲-۱- شاخص‌های ارزیابی «کیفیت فایل داده‌های دریافتی».....	۱۱
۲-۲-۲- راهنمای ارزیابی «کیفیت فایل داده‌های دریافتی».....	۱۲
۲-۳- ارزیابی کیفیت متغیرهای ثبت اداری (q_2).....	۱۵
۲-۳-۱- شاخص‌های ارزیابی «متغیرهای سامانه».....	۱۶
۲-۳-۲- راهنمای ارزیابی «متغیرهای سامانه».....	۱۷
۲-۴- ارزیابی میزان قابل استفاده بودن رکوردهای ثبت اداری (q_3).....	۲۱
۲-۵- محاسبه کیفیت ثبت اداری و میزان قابل استفاده بودن آن ($q_{\alpha 1}$ و $q_{\alpha 2}$).....	۲۲
شاخص‌های ارزیابی در یک نگاه.....	۲۳
منابع.....	۲۵

پیشگفتار

تولید آمارهای با کیفیت، همواره مورد توجه مرکز آمار ایران بوده است. با توجه به رویکرد مرکز آمار ایران برای تغییر شیوه تولید آمار و استفاده بیش‌تر از منابع اداری، خصوصاً انجام سرشماری ثبتی مبنا ۱۴۰۵، این موضوع اهمیت ویژه‌ای پیدا کرده است. به این منظور، موضوع مطالعه و بررسی ارزیابی کیفیت منابع ثبتی و فرایند تولید ثبت‌های آماری بیش از پیش مورد توجه قرار گرفت، به‌طوری که تهیه دستورالعمل ارزیابی کیفیت منابع ثبتی و فرایند تولید آمارهای ثبتی مبنا در دستور کار قرار گرفت. «دستورالعمل ارزیابی کیفیت ثبت اداری و قابلیت تولید ثبت آماری» به منظور پاسخگویی به این نیاز تهیه شده است که در تدوین آن، از تجربیات و نتایج مطالعات ارزنده گذشته استفاده شده است. نسخه حاضر بر اساس دستورالعمل تهیه شده در دفتر استانداردها و نظارت بر طرح‌های آماری در بهمن ۱۳۹۹ آماده شده است و پس از آن، طی ۳ بار ویرایش، بر اساس نظرات دفاتر مرکز، خبرگان و برخی از اعضای کمیته‌ها و کارگروه‌های سرشماری و نیز بر اساس ارزیابی اولیه انجام شده بر روی فایل‌های دریافتی از شرکت پست ج.ا. و وزارت آموزش و پرورش تکمیل و نهایی شده است. در این دستورالعمل دو موضوع «کیفیت ثبت اداری» و «قابلیت تولید ثبت آماری» مورد بررسی قرار می‌گیرد. شاخص‌های مرتبط با هر موضوع به طور خلاصه در جدول ۴ ارائه شده است.

با توجه به این‌که شاخص‌ها و راهنمای تهیه شده، هنوز به طور کامل در ارزیابی کیفیت ثبت‌های اداری مورد استفاده قرار نگرفته است، امکان نیاز به بازنگری شاخص‌ها، فرمول‌ها و راهنمای ارزیابی وجود دارد.

لازم به ذکر است در نسخه فعلی، تمام شاخص‌ها بدون وزن محاسبه می‌شوند. در نسخه بعد، وزن شاخص‌ها بر اساس نتایج بررسی تعداد بیش‌تری از فایل‌های دستگاه‌های اجرایی و میزان کاربردی بودن هر شاخص و نیز دریافت نظر خبرگان، تعیین خواهد شد.

۱- مقدمه

کیفیت آمارهای ثبتی مبنا متأثر از کیفیت داده‌های اداری و روش‌های پردازش مورد استفاده برای تولید ثبت آماری است. کیفیت داده‌های اداری نیز به چگونگی تولید داده‌ها در سیستم اطلاعاتی مرتبط در دستگاه اجرایی بستگی دارد. بنابر این نقطه شروع بررسی کیفیت آمارهای ثبتی، بررسی و تحلیل کیفیت داده‌های ورودی است. این بررسی نشان می‌دهد که آیا ثبت اداری مورد نظر قابلیت استفاده برای تولید آمار را دارد. در این تحلیل، منبع داده باید از جنبه‌های مختلف مورد بررسی و ارزیابی قرار گیرد، مانند میزان دقت داده‌های ثبت شده، زمان مرجع داده‌ها، میزان نزدیکی مفاهیم مورد استفاده در منبع با مفاهیم مورد نیاز، کیفیت ثبت اطلاعات در فرایند اداری مرتبط.

ثبت اداری، اهداف خاص خود را دارد و مسلماً زمانی که بحث تولید آمار از ثبت‌های اداری مطرح است، لازم است ثبت، از جنبه‌های دیگری نیز مورد بررسی قرار گیرد. به عنوان مثال، ممکن است در یک ثبت اداری، از تفکیک یکسانی در کل فرایند تولید داده استفاده شود و برای تولید آمارهای ثبتی دستگاه اجرایی مربوطه، مناسب باشد ولی این تفکیک، مناسب استفاده مرکز آمار ایران نباشد. ممکن است حتی سیستم یکپارچه‌ای در ثبت اداری مورد استفاده قرار گیرد ولی ساختار و نوع فایل در دوره‌های ارسالی مختلف، متفاوت باشد که این موضوع بر میزان قابل استفاده بودن فایل دریافتی برای ایجاد ثبت آماری تأثیرگذار است. از این‌رو در این دستورالعمل علاوه بر بررسی «کیفیت ثبت اداری»، به موضوع «قابلیت تولید ثبت آماری» با استفاده از ثبت اداری مورد نظر نیز پرداخته شده و امتیاز هر بخش به صورت مجزا محاسبه می‌شود.

۲- ارزیابی ثبت اداری و بررسی قابلیت استفاده در تولید ثبت آماری (q_{a1} , q_{a2})

در تولید آمارهای ثبتی مبنا، ثبت اداری باید از جنبه‌های مختلف مورد بررسی قرار گیرد. کیفیت داده‌های تولید شده در ثبت، مناسب بودن فایل ارائه شده با توجه به مواردی نظیر امکان پیوند با سایر ثبت‌ها، وجود رکوردهای تکراری از جمله مواردی است که باید هنگام استفاده از داده‌های اداری مورد توجه قرار گیرد. بر این اساس، چهار حوزه « q_1 »: کیفیت ثبت داده‌ها، « q_2 »: کیفیت فایل داده‌های دریافتی، « q_3 »: کیفیت متغیرهای اصلی ثبت اداری و « q_4 »: ارزیابی میزان قابل استفاده بودن رکوردهای ثبت اداری در این دستورالعمل معرفی شده است که با ترکیبی از شاخص‌های ارائه شده در هر حوزه می‌توان دو موضوع زیر را بررسی کرد:

۱. ارزیابی کیفیت ثبت اداری (q_{a1})

۲. ارزیابی قابلیت تولید ثبت آماری با استفاده از ثبت اداری مورد نظر (q_{a2})

شاخص‌های مرتبط با هر یک از موضوعات مطرح شده در بالا در جدول ۴ ارائه شده است.

۲-۱- کیفیت ثبت داده‌ها (q.)

شاخص q. بر اساس اطلاعات به‌دست آمده از شناخت فرایند تولید داده‌های اداری و با استفاده از اطلاعات جدول ۱ به‌دست می‌آید. در این جدول «کیفیت ثبت داده‌ها» مورد بررسی و ارزیابی قرار می‌گیرد. شاخص q. میانگین موزون امتیازهای مندرج در جدول ۱ است.

$$q. = \frac{\sum_{i=1}^m w_i x_i}{\sum_{i=1}^m w_i}$$

x_i امتیاز کسب شده ثبت اداری در شاخص α_m

w_i وزن شاخص α_m

m تعداد شاخص‌هایی که برای ثبت اداری مورد نظر محاسبه شده است.

۲-۱-۱- شاخص‌های ارزیابی «کیفیت ثبت داده‌ها»

شاخص‌های ارزیابی «کیفیت ثبت داده‌ها» در جدول ۱ ارائه شده است. ارزیابی تمام شاخص‌های این جدول، پس از مذاکره با کارشناس/کارشناسان مطلع از فرایند تولید داده‌های مورد بررسی تکمیل می‌شود. برای تکمیل این اطلاع، حضور در سازمان، مذاکره با کارشناس/کارشناسان مرتبط و برگزاری چندین جلسه با آن‌ها ضروری است.

جدول ۱- ارزیابی «کیفیت ثبت داده‌ها»

ردیف	عنوان شاخص	راهنمای ارزیابی	وزن شاخص (w)	امتیاز کسب شده (x)
۱	۲	۳	۴	۵
۱	آیا عنوان متغیر، تمامی رخدادها را پوشش می‌دهد؟	- میزان پوشش ۱۰۰٪ (۱ امتیاز) - پوشش کمتر از ۱۰۰ و بیشتر از ۹۰٪ (۰,۸ امتیاز) - پوشش کمتر از ۹۰ و بیشتر از ۷۵٪ (۰,۶ امتیاز) - کمتر از ۷۵ و بیشتر از ۵۰٪ (۰,۴ امتیاز) - کمتر از ۵۰٪ (۰ امتیاز)	—	
۲	آیا تعاریف و مفاهیم، کامل و مقایسه‌پذیر است؟	- تمام تعاریف و مفاهیم استاندارد است (۱ امتیاز) - برخی از تعاریف و مفاهیم، استاندارد و برخی درون دستگاه، یکسان است (۰,۷ امتیاز) - تمام تعاریف و مفاهیم، فقط در داخل دستگاه یکسان است. (۰,۳ امتیاز) - تعاریف و مفاهیم، یکسان نیست (۰ امتیاز)	—	
۳	آیا استانداردها و قراردادهای مشترک در مورد زمان (از قبیل سری زمانی، بازه زمانی و مقطع زمانی) برای ثبت متغیرهای مورد نظر رعایت می‌شود؟	- بله (۱ امتیاز) - تا حدودی (۰,۵ امتیاز) - خیر (۰ امتیاز)	—	

ردیف	عنوان شاخص	راهنمای ارزیابی	وزن شاخص (w)	امتیاز کسب شده (x)
۱	۲	۳	۴	۵
۴	آیا استانداردها و قراردادهای مشترک در مورد مشخصات مکانی (از قبیل سطوح جغرافیایی، یکپارچگی داده‌ها، ساختار و قابلیت استفاده در سیستم‌های اطلاعات جغرافیایی GIS و...) برای ثبت متغیرهای مورد نظر رعایت می‌شود؟	- بله (۱ امتیاز) - تا حدودی (۰,۵ امتیاز) - خیر (۰ امتیاز)	—	
۵	آیا استانداردها و قراردادهای مشترک در مورد تفکیک‌ها (از قبیل سن، مدرک تحصیلی و...) و طبقه‌بندی‌های بین‌المللی (از قبیل I S I C, I S C E D و...) برای ثبت قلم آماری مذکور رعایت می‌شود؟	- بله (۱ امتیاز) - تا حدودی (۰,۵ امتیاز) - خیر (۰ امتیاز)	—	
۶	آیا داده‌ها در پایین‌ترین نقطه، ثبت می‌شوند؟	- بله (۱ امتیاز) - یک ایستگاه بالاتر از پایین‌ترین نقطه، ثبت می‌شوند (۰,۸ امتیاز) - دو ایستگاه بالاتر از پایین‌ترین نقطه ثبت می‌شوند (۰,۶ امتیاز) - سه ایستگاه بالاتر از پایین‌ترین نقطه ثبت می‌شوند (۰,۴ امتیاز) - بیش از سه ایستگاه بالاتر از پایین‌ترین نقطه ثبت می‌شوند (۰,۲ امتیاز)	—	
۷	آیا داده‌ها به هنگام وقوع عملیات ثبت می‌شوند؟	- بله (۱ امتیاز) - داده‌ها یک روز بعد از وقوع عملیات، ثبت می‌شوند (۰,۸ امتیاز) - داده‌ها دو روز بعد از وقوع عملیات، ثبت می‌شوند (۰,۶ امتیاز) - داده‌ها سه روز بعد از وقوع عملیات، ثبت می‌شوند (۰,۴ امتیاز) - داده‌ها بیش از سه روز بعد از وقوع عملیات، ثبت می‌شوند (۰,۲ امتیاز)	—	
۸	آیا ورود اطلاعات به سیستم همزمان با انجام عملیات ثبت صورت می‌گیرد؟	- بله (۱ امتیاز) - ورود اطلاعات یک هفته بعد از وقوع عملیات، انجام می‌شود (۰,۷ امتیاز) - ورود اطلاعات دو هفته بعد از وقوع عملیات، انجام می‌شود (۰,۴ امتیاز) - ورود اطلاعات بیش از دو هفته بعد از وقوع عملیات، انجام می‌شود (۰,۱ امتیاز)	—	

ردیف	عنوان شاخص	راهنمای ارزیابی	وزن شاخص (w)	امتیاز کسب شده (x)
۱	۲	۳	۴	۵
۹	آیا عمل ثبت در فرآیند تولید توسط افراد مسئول و پاسخگو انجام می شود؟	- در کل فرایند وجود دارد (۱ امتیاز) - در برخی از مراحل فرایند وجود دارد (۰.۵ امتیاز) - در هیچ مرحله ای وجود ندارد (۰ امتیاز)	—	
۱۰	آیا فرم های عملیاتی که در فرآیند تولید اقلام آماری به کار می روند برای تأمین اطلاعات مورد نیاز طراحی مناسبی دارند؟	- بله (۱ امتیاز) - تا حدودی (۰.۵ امتیاز) - خیر (۰ امتیاز)	—	
۱۱	آیا برای کنترل و تأیید صحت ثبت در مراحل مختلف، مکانیزمی وجود دارد؟	- در تمام مراحل وجود دارد (۱ امتیاز) - در برخی از مراحل وجود دارد (۰.۵ امتیاز) - در هیچ مرحله ای وجود ندارد (۰ امتیاز)	—	
۱۲	آیا الزام قانونی برای ثبت متغیر وجود دارد؟	- بله (۱ امتیاز) - خیر (۰ امتیاز)	—	
۱۳	آیا عمل انتقال در فرآیند تولید توسط افراد مسئول و پاسخگو انجام می شود؟	- در کل فرایند وجود دارد (۱ امتیاز) - در برخی از مراحل فرایند وجود دارد (۰.۵ امتیاز) - در هیچ مرحله ای وجود ندارد (۰ امتیاز)	—	
۱۴	آیا داده های ثبتی تولید شده از مسیرهای مناسب به سطوح بالاتر منتقل می شوند؟	- بله (۱ امتیاز) - خیر (۰ امتیاز)	—	
۱۵	آیا در مسیر انتقال، هر ایستگاه و وظایف خود (از قبیل کنترل، تجمع درست یا اقدام به موقع برای انتقال اطلاعات و ...) را به خوبی انجام می دهد؟	- تمام ایستگاه ها وظایف خود را به خوبی انجام می دهند (۱ امتیاز) - برخی از ایستگاه ها وظایف خود را به خوبی انجام می دهند (۰.۵ امتیاز) - هیچ کدام از ایستگاه ها وظایف خود را به خوبی انجام نمی دهند (۰ امتیاز)	—	
۱۶	آیا تکنولوژی انتقال، متناسب با اهمیت و فرآیند تولید داده ها است؟	- بله (۱ امتیاز) - خیر (۰ امتیاز)	—	
۱۷	آیا امنیت داده ها و اطلاعات هنگام انتقال رعایت می شود؟	- بله (۱ امتیاز) - خیر (۰ امتیاز)	—	

۲-۱-۲- راهنمای ارزیابی «کیفیت ثبت داده ها»

• شاخص ۱- آیا عنوان متغیر، تمامی رخدادهای مربوطه را پوشش می دهد؟

ارزیابی این شاخص با توجه به میزان پوشش رخدادهای انجام می شود.

- شاخص ۲- آیا تعاریف و مفاهیم، کامل و مقایسه‌پذیر است؟

مقایسه‌پذیر بودن تعاریف دارای اهمیت زیادی است. ممکن است در تمام فرایند تولید داده‌های ثبتی، تعاریف تمام متغیرها توسط مرکز آمار استاندارد شده باشد، در این صورت امتیاز کامل درج می‌شود. اما ممکن است تعاریف برخی از متغیرها استاندارد شده باشد ولی برای برخی دیگر، علاوه بر نبود تعریف استاندارد، از تعاریف یکسان در فرایند تولید داده‌ها استفاده می‌شود. در این صورت امتیاز ۰,۷ درج می‌شود. در صورتی که هیچ یک از متغیرها توسط مرکز آمار ایران استاندارد نشده است ولی در تمام فرایند تولید داده‌ها، از تعاریف یکسان استفاده می‌شود، امتیاز ۰,۳ درج می‌شود. در صورتی که تعاریف یکسان مورد استفاده قرار نمی‌گیرد، هیچ امتیازی به این شاخص اختصاص داده نمی‌شود.

- شاخص ۳- آیا استانداردها و قراردادهای مشترک در مورد زمان (از قبیل سری زمانی، بازه زمانی و

مقطع زمانی) برای ثبت متغیرهای مورد نظر رعایت می‌شود؟

ارزیابی با توجه به میزان به کارگیری قراردادهای مشترک زمانی، انجام می‌شود.

- شاخص ۴- آیا استانداردها و قراردادهای مشترک در مورد مشخصات مکانی (از قبیل سطوح

جغرافیایی، یکپارچگی داده‌ها، ساختار و قابلیت استفاده در سیستم‌های اطلاعات جغرافیایی GIS و...)

برای ثبت متغیرهای مورد نظر رعایت می‌شود؟

ارزیابی با توجه به میزان به کارگیری قراردادهای مشترک در مورد مشخصات مکانی، انجام می‌شود.

- شاخص ۵- آیا استانداردها و قراردادهای مشترک در مورد تفکیک‌ها (از قبیل سن، مدرک تحصیلی

و ...) و طبقه‌بندی‌های بین‌المللی (از قبیل ISCED, ISIC و...) برای ثبت قلم آماری مذکور رعایت

می‌شود؟

ارزیابی با توجه به میزان به کارگیری قراردادهای مشترک در مورد مشخصات مکانی، انجام می‌شود.

- شاخص ۶- آیا داده‌ها در پایین‌ترین نقطه، ثبت می‌شوند؟

منظور از پایین‌ترین نقطه، معمولاً جایی است که عملیات مربوط به آن متغیر انجام می‌شود. پایین‌ترین نقطه تولید داده‌های یک متغیر بنا به تعریف و هدف آن می‌تواند تغییر یابد. برای مثال چنانچه آمار روزانه تولید لامپ یک کارخانه موردنظر باشد، انتهای خط تولید و واحد کنترل کیفیت هریک می‌توانند بنابه هدفی، به‌عنوان پایین‌ترین نقطه تولید تعریف شود. نقطه انتهای خط تولید می‌تواند داده‌های خام تولید لامپ را ثبت کند. این اطلاع می‌تواند مثلاً برای آگاهی از مواد اولیه مصرف شده در یک روز مورد استفاده قرار گیرد. این در حالی است که نقطه واحد

کنترل کیفیت می‌تواند موجودی انبار و یا درآمد روزانه حاصل از تولید را ثبت کند. به عنوان مثالی دیگر، فرایند ثبت در شماره‌گذاری خودرو در پایین‌ترین نقطه یعنی دقیقاً همانجایی که عمل شماره‌گذاری انجام می‌شود، صورت می‌گیرد.

• شاخص ۷- آیا داده‌ها به هنگام وقوع عملیات، ثبت می‌شوند؟

ثبت داده‌ها باید هم زمان با وقوع عملیات انجام شود زیرا در غیر این صورت ممکن است ثبت آن به فراموشی سپرده شود و یا با انباشته شدن داده‌هایی که باید ثبت شوند، خطر کاهش دقت پیش می‌آید. ارزیابی شاخص با توجه به فاصله زمانی بین وقوع عملیات تا ثبت آن (بر حسب روز) انجام می‌شود.

• شاخص ۸- آیا ورود اطلاعات به سیستم همزمان با انجام عملیات ثبت صورت می‌گیرد؟

در صورت استفاده از فرم‌های کاغذی برای دریافت اطلاعات، باید فعالیت دیگری به منظور ورود اطلاعات به سیستم انجام شود. ورود اطلاعات ممکن است با خطا همراه باشد. بهترین حالت، ورود اطلاعات به سیستم، همزمان با انجام عملیات ثبت است. در این صورت علاوه بر افزایش دقت در اطلاعات وارد شده، امکان تولید داده‌های به‌هنگام نیز فراهم خواهد شد. ارزیابی شاخص با توجه به فاصله زمانی بین ورود اطلاعات به سیستم تا ثبت آن (بر حسب هفته) انجام می‌شود.

• شاخص ۹- آیا عمل ثبت در فرآیند تولید توسط افراد مسئول و پاسخگو انجام می‌شود؟

در فرم‌های عملیاتی باید مکانی برای تأیید شخص مسئول آن عملیات، پیش‌بینی شود که هم داده‌ها را کنترل نمایند و هم مسئولیت درستی آن‌ها را بر عهده گیرند. به این ترتیب بروز هر نوع مشکلی در ثبت داده‌ها قابل پیگیری خواهد بود. در صورتی که از فرم استفاده می‌شود، می‌توان از امضای شخص پایین فرم استفاده کرد یا اگر سیستمی در کار است، نام کاربری و کلمه عبور مسئول این کار است. به عنوان مثال وقتی در بانک، مراجعه کننده زیر برگه یا فیش بانکی را امضا می‌نماید، صحت اطلاعات مندرج آن‌را تأیید می‌کند. پس از آن نیز وقتی متصدی بعد از عملیات بانکی آن‌را مهر و امضا کرده و صحت اطلاعات را تأیید می‌کند.

در صورتی که در تمام مراحل مختلف ثبت، افراد مسئول و پاسخگو وجود دارد، گزینه «بله» انتخاب می‌شود. در صورتی که در هیچ مرحله‌ای فرد مسئول و پاسخگو وجود نداشته باشد، گزینه «خیر» و در سایر موارد، گزینه «تا حدودی» انتخاب می‌شود.

• شاخص ۱۰- آیا فرم‌های عملیاتی که در فرآیند تولید اقلام آماری به کار می‌روند برای تأمین اطلاعات

مورد نیاز طراحی مناسبی دارند؟

ارزیابی با توجه به میزان مناسب بودن طراحی فرم‌ها برای جمع‌آوری داده‌ها و اطلاعات مورد نیاز، انجام می‌شود.

• شاخص ۱۱- آیا برای کنترل و تأیید صحت ثبت در مراحل مختلف، مکانیزمی وجود دارد؟

برخی از مکانیزم‌های موجود برای تأیید صحت ثبت، وجود بار مالی یا مدعیان درون سازمانی یا برون سازمانی است. به عنوان نمونه، در مورد آمار «تعداد معتادین تحت بازپروری»، فعالیت‌های زیر به‌طور خلاصه اتفاق می‌افتد:

۱- دستگیری معنادین توسط مقامات انتظامی یا تحویل آنها توسط خانواده‌ها.

۲- اعزام به مراکز بهزیستی.

۳- تحویل مرکز شده و در دفتر ثبت می‌شود.

۴- درمان انجام می‌شود.

۵- معناد ترخیص می‌شود.

در این مجموعه فعالیت‌ها، از مرحله سوم می‌توان به عنوان یک نقطه مناسب برای ثبت اطلاعات بهره برد. زیرا در این مرحله امکان بررسی صحت اطلاعات میسر دارد. وجود مدعی برون سازمانی (نیروی انتظامی) و درون سازمانی (سایر واحدهای بهزیستی) امکان بررسی صحت این اطلاع را فراهم می‌کند.

در صورتی که در فرایند ثبت و انتقال داده‌ها، سیستم وجود داشته باشد، مکانیزم کنترل می‌تواند مواردی نظیر بررسی ارتباط منطقی بین متغیرها یا تعیین محدوده قابل قبول برای هر متغیر باشد و با اعلام هشدار یا خطا، اپراتور را از وجود خطای احتمالی آگاه سازد. به عنوان نمونه در صورت درج مقدار ۱۰ سال سابقه بیمه یا کار برای یک فرد ۱۸ ساله با مدرک لیسانس، باید در سیستم، هشدار داده شود. یا در صورتی که تاریخ ازدواج کمتر از تاریخ تولد باشد، پیام خطا ظاهر شود.

ارزیابی این شاخص، با توجه به میزان تایید و کنترل موجود در فرایند صورت می‌گیرد.

• شاخص ۱۲- آیا الزام قانونی برای متغیر وجود دارد؟

در صورت وجود الزام قانونی برای ثبت اطلاعات، تکمیل‌کننده فرم یا پاسخگو موظف می‌شود که علاوه بر ثبت اطلاع، اطلاع صحیح ارائه کند. از این رو از یک سو باعث کاهش بی‌پاسخی متغیر و از سوی دیگر باعث افزایش دقت اطلاعات درج شده می‌شود. در صورت وجود الزام قانونی، گزینه «بله» و در غیر این صورت، گزینه «خیر» انتخاب می‌شود.

• شاخص ۱۳- آیا عمل انتقال در فرآیند تولید توسط افراد مسئول و پاسخگو انجام می‌شود؟

در انتقال داده‌ها به سطوح بالاتر، باید جایگاهی برای افراد مسئول پیش‌بینی شود که هم داده‌ها را کنترل نمایند و هم مسئولیت درستی آن‌ها را بر عهده گیرند. به این ترتیب بروز هر نوع مشکلی در انتقال داده‌ها قابل پیگیری خواهد بود. در صورتی که در تمام مراحل مختلف ثبت، افراد مسئول و پاسخگو وجود دارد، گزینه «بله» انتخاب می‌شود. در صورتی که در هیچ مرحله‌ای فرد مسئول و پاسخگو وجود نداشته باشد، گزینه «خیر» و در سایر موارد، گزینه «تا حدودی» انتخاب می‌شود.

• شاخص ۱۴- آیا داده‌های ثبتی تولید شده از مسیرهای مناسب به سطوح بالاتر منتقل می‌شوند؟

داده‌ای که در طی یک فرایندی تولید، ثبت می‌شود باید برای طی مراحل دیگر (مثلاً کنترل، تجمیع) به مراحل و سطوح بالاتر منتقل شود. در صورت مناسب بودن مسیر، گزینه «بله» و در غیر این صورت گزینه «خیر» انتخاب می‌شود.

- شاخص ۱۵- آیا در مسیر انتقال، هر ایستگاه وظایف خود (از قبیل کنترل، جمع‌آوری یا اقدام به موقع برای انتقال اطلاعات و ...) را به خوبی انجام می‌دهد؟

برای هر ایستگاه کاری، وظایف تعریف شده‌ای وجود دارد. به عنوان مثال، مرکز استان موظف است داده‌های شهرستان‌ها را جمع‌آوری و پس از جمع‌آوری به اداره مرکزی ارسال کند. اداره مرکزی هم موظف است آن‌ها را کنترل و ارائه کند. این وظایف تعریف شده باید به نحو احسن انجام شود، زیرا کوچک‌ترین کم‌کاری در ایستگاه پایین بزرگ‌ترین خطاها را در مرکز خواهد داشت.

ارسال به موقع نیز جزو وظایفی است که باید با دقت بالا انجام شود. فرض کنید قرار است داده‌های تولیدی هر شهرستان هر ماه به اداره مرکزی ارسال شود. اگر یک شهرستان در ارسال داده‌ها تاخیر داشته باشد، این امر در به روز بودن آمار آن سازمان خلل وارد می‌کند.

- شاخص ۱۶- آیا تکنولوژی انتقال، متناسب با اهمیت و فرآیند تولید داده‌ها است؟

تکنولوژی انتقال، باید متناسب با اهمیت و فرآیند تولید داده‌ها باشد. تکنولوژی صرفاً به معنی استفاده از امکانات اینترنتی نیست. بلکه در برخی موارد می‌توان از تکنولوژی‌های ساده نیز استفاده کرد ولی حتماً باید متناسب با اهمیت متغیر مورد نظر باشد. اگر نیاز است آماری بصورت لحظه‌ای به روز شود باید از تکنولوژی تحت وب و آنلاین استفاده کند. به عنوان نمونه سیستم کارت هوشمند سوخت به دلیل حساسیت زیاد آن از تکنولوژی آنلاین منطقه‌ای استفاده می‌کند که هر ۲ ساعت، داده‌های مرکز تهران را به روز می‌کند. انتقال می‌تواند به روش‌های زیر انجام شود:

- ✓ استفاده از امکانات اینترنت و یا به صورت بر خط
- ✓ با استفاده از شبکه‌های محلی یا FTP انجام
- ✓ انتقال با استفاده از فلاپی یا لوح فشرده
- ✓ انتقال با استفاده از کاغذ یا فرم.

با توجه به اهمیت فرایند تولید داده‌ها، ارزیابی به صورت «بله» و «خیر» انجام می‌شود.

- شاخص ۱۷- آیا امنیت داده‌ها و اطلاعات هنگام انتقال رعایت می‌شود؟

- مسیر و تکنولوژی انتقال داده‌ها باید امن باشد به گونه‌ای که داده‌ها و اطلاعات، مخدوش نشوند. در صورت امن بودن انتقال داده‌ها، گزینه «بله» و در غیر این صورت گزینه «خیر» انتخاب می‌شود.

۲-۲- کیفیت فایل داده‌های دریافتی (q1)

شاخص q1 بر اساس اطلاعات به دست آمده از شناخت ثبت اداری و با استفاده از اطلاعات جدول ۱ به دست می‌آید. در این جدول «مشخصات فنی ثبت اداری» مورد بررسی و ارزیابی قرار می‌گیرد. شاخص q1، میانگین موزون امتیازهای حاصل از «مشخصات فنی ثبت اداری» است.

$$q_1 = \frac{\sum_{i=1}^r w e_i y_i}{\sum_{i=1}^r w e_i}$$

y_i امتیاز کسب شده ثبت اداری در شاخص α_i

$w e_i$ وزن شاخص α_i

r تعداد شاخص‌هایی که برای ثبت اداری مورد نظر محاسبه شده است.

۲-۲-۱- شاخص‌های ارزیابی «کیفیت فایل داده‌های دریافتی»

شاخص‌های ارزیابی «مشخصات فنی ثبت اداری» در جدول ۲ ارائه شده است. ارزیابی تمام شاخص‌های این جدول، پس از دریافت فایل داده از دستگاه اجرایی تکمیل می‌شود.

جدول ۲- ارزیابی «کیفیت فایل داده‌های دریافتی»

ردیف	عنوان شاخص	راهنمای ارزیابی	وزن شاخص ($w e$)	امتیاز کسب شده (y)
۱	آیا فایل‌های داده ارسالی با موضوع یکسان، یکپارچه است؟	- بله (۱ امتیاز) - خیر ولی ساختار فایل‌ها یکسان و قابل یکپارچه‌سازی است (۰,۵ امتیاز) - خیر و ساختار فایل‌ها یکسان نبوده و قابل یکپارچه‌سازی نیست (۰ امتیاز)	—	۵
۲	آیا مدل داده برای فایل داده در اختیار مرکز قرار گرفته است؟ توضیح: برای هر موجودیت باید یک دیتا مدل ارائه شود	- بله (۱ امتیاز) - تا حدودی (۰,۵ امتیاز) - خیر (۰ امتیاز)	—	۵
۳	آیا فراداده‌های مورد نیاز، همراه با داده‌ها در اختیار مرکز قرار گرفته است؟	- به‌صورت کامل در اختیار مرکز قرار گرفته است (۱ امتیاز) - به‌صورت ناقص در اختیار مرکز قرار گرفته است (۰,۵ امتیاز) - در اختیار مرکز قرار نگرفته است (۰ امتیاز)	—	۵
۴	آیا موارد اصلاحی مورد نظر مرکز (نظیر ساختار، محتوا و نوع فایل)، در دوره‌های بعدی اعمال شده است؟	- بله (۱ امتیاز) - خیر (۰ امتیاز)	—	۵
۵	آیا ساختار و نوع فایل، در دوره‌های مختلف ارسال یکسان است؟	- هم ساختار و هم نوع فایل یکسان است (۱ امتیاز) - فقط ساختار یا فقط نوع فایل یکسان است (۰,۵ امتیاز) - ساختار و نوع فایل، یکسان نیستند (۰ امتیاز)	—	۵

ردیف	عنوان شاخص	راهنمای ارزیابی	وزن شاخص (we)	امتیاز کسب شده (y)
۱	۲	۳	۴	۵
۶	آیا هر موجودیت در فایل داده ارسالی، دارای متغیر/ متغیرهای یکتاکننده است؟	- بله (۱ امتیاز) - خیر (۰ امتیاز)	—	
۷	آیا فایل داده دارای متغیر/ متغیرهای اطلاعاتی با قابلیت پیوند با فایل‌های داده سایر دستگاه‌های اجرایی است؟	- بله (۱ امتیاز) - خیر (۰ امتیاز)	—	
۸	آیا ثبت اداری قابلیت پیوند با اطلاعات مکانی برای تفکیک‌های جغرافیایی را دارا است؟	- اطلاعات در سطح پارسل قابل تولید است (۱ امتیاز) - اطلاعات در سطح بلوک قابل تولید است (۰,۹ امتیاز) - اطلاعات در سطح شهر/آبادی قابل تولید است (۰,۷۵ امتیاز) - اطلاعات در سطح دهستان قابل تولید است (۰,۶ امتیاز) - اطلاعات در سطح شهرستان قابل تولید است (۰,۴۵ امتیاز) - اطلاعات در سطح استان قابل تولید است (۰,۳ امتیاز) - اطلاعات در سطح کل کشور قابل تولید است (۰,۱۵ امتیاز) - چنین امکانی وجود ندارد (۰ امتیاز)	—	
۹	آیا داده‌ها بر اساس زمان‌بندی مورد توافق، دریافت شده است؟	- بله (۱ امتیاز) - خیر (۰ امتیاز)	—	

۲-۲-۲- راهنمای ارزیابی «کیفیت فایل داده‌های دریافتی»

• شاخص ۱- آیا فایل‌های داده ارسالی با موضوع یکسان، یکپارچه هستند؟

فایل داده یکپارچه، فایلی است که تمام داده‌های مربوط به یک موجودیت^۱ در تمام سطوح را در یک جدول در بر داشته باشد. به‌عنوان نمونه اگر فایل مربوط به دانش‌آموزان، تمام اطلاعات مرتبط را برای تمام نقاط کشور در یک جدول ارائه دهد، این فایل، یک فایل یکپارچه است که در این صورت امتیاز ۱ اختصاص می‌یابد. ممکن است تمام

^۱ موجودیت (Entity): مفهوم کلی پدیده، شیء یا فرد است که در مورد آن می‌خواهیم اطلاع داشته باشیم. هر نوع موجودیت دارای مجموعه‌ای از صفات خاصه (attribute) است که ویژگی جداکننده یک نوع موجودیت از نوع دیگر محسوب می‌شود. به عنوان مثال، موجودیت «دانشجو» در سیستم دانشگاه می‌تواند دارای صفات خاصه: نام و نام خانوادگی، سن، سال تولد، رشته تحصیلی، مقطع تحصیلی، سال ورود به دانشگاه و ... باشد. موجودیت دیگر در این سیستم «درس» است که دارای صفات خاصه: کد درس، نام درس، تعداد واحد و ... است.

داده‌ها در یک جدول نباشد ولی ساختار جداول مختلف به گونه‌ای است که امکان یکپارچه کردن آن به سادگی وجود دارد. به عنوان مثال اگر فایل دانش آموزان در ۳۱ جدول مربوط به ۳۱ استان ارائه شود ولی ساختار فایل‌ها یکسان باشد، در این صورت امکان یکپارچه کردن فایل به سادگی وجود دارد و امتیاز ۱ تعلق می‌گیرد. به عنوان مثالی دیگر، اگر تعدادی از فیلدهای مرتبط با دانش آموزان در یک جدول و برخی در جدول دیگر باشد ولی رکوردها به صورتی باشد که بتوان دو جدول را به راحتی ترکیب نمود، در این صورت می‌توان گفت فایل یکپارچه است و امتیاز ۱ تعلق می‌گیرد. در مواردی که امکان یکپارچه کردن فایل‌ها وجود ندارد یا یکپارچه کردن آن‌ها بسیار زمان‌بر بوده و نیاز به یکسان‌سازی‌های بسیار روی جداول مختلف است، در این صورت، امتیاز ۰ تعلق می‌گیرد. در مواردی که یکپارچه کردن زیاد زمان‌بر نبوده و کار کارشناسی زیادی لازم نداشته باشد، در این صورت امتیاز ۰.۵ تعلق می‌گیرد.

شاخص ۲- آیا مدل داده برای فایل داده در اختیار مرکز قرار گرفته است؟

به ازای هر موجودیت در فایل داده باید مدل داده مربوط تهیه شود. مدل داده شامل اطلاعات جداول زیر است:

مدل داده	
نام دستگاه اجرایی	
نام فایل	
(موضوع فایل+تاریخ ارسال+علامت اختصاری دستگاه)	
نام موجودیت	

نام لاتین فیلد	نام فارسی فیلد	نوع فیلد	سایز فیلد	کلید (بلی / نه)	فرمت فیلد	نام موجودیت وابسته	نام فیلد موجودیت وابسته

در صورتی که مدل داده کاملاً منطبق با جداول بالا است، گزینه «بله» انتخاب می‌شود. در صورتی که تنها بخشی از اطلاع را در بر دارد، گزینه «تأخودی» و در صورتی که مدل داده ارسال نشده یا هیچ تطابقی با فرمت جدول بالا نداشته باشد، گزینه «خیر» انتخاب می‌شود.

• شاخص ۳- آیا فراداده‌های مورد نیاز، همراه با داده‌ها در اختیار مرکز قرار گرفته است؟

فراداده‌های مورد نیاز مرکز برای استفاده درست از داده‌ها شامل موارد زیر است:

- ✓ تعاریف و مفاهیم به کار رفته (در صورت یکپارچه نبودن سیستم، چنانچه تعاریف و مفاهیم به کار رفته در واحدهای مختلف جمع‌آوری داده‌ها (مانند استان) متفاوت است، ذکر تعاریف همه واحدها الزامی است)
- ✓ طبقه‌بندی‌های به کار رفته

- ✓ کدها و توصیف‌های مرتبط (به‌عنوان نمونه فرض شود که در متغیر جنسیت از کد ۰ برای توصیف «زن» و از کد ۱ برای توصیف «مرد» استفاده شده است. این توضیح برای این بند لازم است آورده شود).
- ✓ جامعه تحت پوشش ثبت (به عنوان مثال باید مشخص شود که نمونه فایل دریافتی از آموزش و پرورش شامل محصلین است یا فارغ التحصیلان آموزش و پرورش را شامل نیز شامل می‌شود)
- ✓ زمان مرجع اطلاعات (داده‌های ارسالی مربوط به چه زمانی است. به‌عنوان نمونه باید مشخص شود که فایل دریافتی از آموزش و پرورش، محصلین و فارغ التحصیلان چه سال‌هایی را شامل می‌شود)
- ✓ مشخصات دوره‌های به‌هنگام‌سازی فایل (آخرین زمان به‌هنگام سازی داده‌ها، چه زمانی بوده است، امکان به‌هنگام سازی در چه دوره‌ها یا زمان‌هایی امکان‌پذیر است)
- ✓ دستورالعمل‌های مربوط به بررسی، اصلاح، ارتباط بین متغیرها، مقادیر مجاز یا خارج از محدوده
- ✓ نام و نام خانوادگی و شماره تلفن کارشناس برای برقراری ارتباط

• شاخص ۴- آیا موارد اصلاحی مورد نظر مرکز (نظیر ساختار، محتوا و نوع فایل)، در دوره‌های بعدی اعمال شده است؟

چنانچه مرکز آمار خواستار اصلاح فایل از نظر ساختار، محتوا، نوع فایل و ... باشد، در صورتی که در دوره‌های بعدی، اصلاحات مورد نظر اعمال شده است، گزینه «بله» و در غیر این صورت گزینه «خیر» انتخاب می‌شود.
تذکر: برای اولین ارسال یا در مواردی که اصلاحی مورد نظر مرکز نبوده است، نمره‌ای منظور نمی‌شود (در ارزیابی سؤال، علامت - درج می‌شود).

• شاخص ۵- آیا ساختار و نوع فایل، در دوره‌های مختلف ارسال یکسان است؟

منظور از ساختار فایل، تعداد فیلدها، مکان قرار گرفتن فیلدها در فایل، نوع فیلد، طول فیلد و ... است. منظور از نوع فایل، دریافت فایل با فرمت‌های Oracle, SQL, Accssess و ... با مشخصات یکسان در دوره‌های مختلف است.
تذکر: در صورتی که مرکز آمار خواستار اصلاح ساختار و نوع فایل بوده است، این سؤال تکمیل نشده و در ارزیابی محسوب نمی‌شود (در ارزیابی سؤال، علامت - درج می‌شود).

• شاخص ۶- آیا هر موجودیت در فایل داده ارسالی، دارای متغیر / متغیرهای یکتاکننده است؟

وجود فیلدهای یکتاکننده تحت عنوان Primary Key در فایل داده مورد نظر است.

• شاخص ۷- آیا فایل داده دارای متغیر / متغیرهای اطلاعاتی با قابلیت پیوند با فایل‌های داده سایر دستگاه‌های اجرایی است؟

مهمترین متغیر قابل پیوند، «شناسه ملی» برای پیوند افراد و «کد پستی» برای پیوند مکان‌ها است. اما در برخی از فایل‌ها، این کدها یا اصلاً وجود ندارد یا به دلایلی مانند اجباری نبودن ثبت آن‌ها یا قدیمی بودن رکوردها، این اطلاع در فایل ناقص است. در این صورت از اطلاعات کمکی برای پیوند رکوردها استفاده می‌شود. در صورت وجود و کافی بودن این متغیرها، گزینه «بله» و در غیر این صورت گزینه «خیر» انتخاب می‌شود. به‌عنوان نمونه متغیرهای «نام و نام خانوادگی» و «سن» به تنهایی برای پیوند رکوردها کافی نیستند. اطلاعات دیگری نظیر نام پدر و کد پستی نیز باید وجود داشته باشد که بتوان این پیوند را به درستی ایجاد کرد.

• **شاخص ۸- آیا ثبت اداری قابلیت پیوند با اطلاعات مکانی برای تفکیک‌های جغرافیایی را دارا است؟**
تفکیک‌ها در پایین‌ترین سطح جغرافیایی (پارسل) مورد نظر است. چنانچه فیلد کد پستی یا فیلدهای اطلاعاتی دیگری در فایل وجود داشته باشد، می‌توان تفکیک‌های جغرافیایی را به دست آورد. در این شاخص تنها به وجود فیلد کد پستی یا فیلد مشابه پرداخته می‌شود. بی‌پاسخی این فیلد در بخش بعد مورد ارزیابی قرار می‌گیرد.

• **شاخص ۹- آیا داده‌ها بر اساس زمان‌بندی مورد توافق، دریافت شده است؟**
با توجه به زمان‌بندی مورد توافق برای دریافت داده‌ها، چنانچه فایل داده‌ها در زمان مورد توافق یا زودتر از آن در اختیار مرکز آمار قرار گرفت، امتیاز ۱ و در غیر این صورت امتیاز ۰ منظور می‌شود.
دقت شود که در ایت ضاخص، تنها موضوع زمان مطرح است و کیفیت فایل دریافت شده، در سایر شاخص‌های این بخش مورد بررسی قرار می‌گیرد.

۲-۳- ارزیابی کیفیت متغیرهای ثبت اداری (q_۲)

شاخص‌های ارزیابی «متغیرهای ثبت اداری» در جدول ۲ ارائه شده است. برای ارزیابی متغیرهای ثبت اداری توجه به موضوعات زیر ضروری است:

- ارزیابی شاخص‌های ۱ تا ۴ با توجه به شناخت ثبت اداری و فراداده‌های دریافت شده، انجام می‌شود.
- شاخص‌های ۵ تا ۹ پس از دریافت فایل داده قابل محاسبه است.
- در این جدول، هر متغیر ثبت اداری به‌صورت مجزا مورد ارزیابی قرار گرفته و میانگین موزون کسب شده برای هر متغیر محاسبه می‌شود (q_{۲i}). سپس مقدار شاخص q_۲، بر اساس میانگین حسابی امتیاز تمام متغیرها، به دست می‌آید.

نکته: از آن‌جا که نمره کیفیت هر شاخص به طور مجزا محاسبه می‌شود، می‌توان از آن به‌عنوان معیاری برای انتخاب متغیر بهتر در جانهی بی‌پاسخی پس از پیوند چند ثبت استفاده نمود.

در این جدول، هر متغیر ثبت اداری به صورت مجزا مورد ارزیابی قرار می‌گیرد. برای محاسبه شاخص q_2 ، ابتدا میانگین موزون کسب شده برای هر متغیر محاسبه شده (q_{2i}) و سپس مقدار شاخص q_2 بر اساس میانگین حسابی امتیاز تمام متغیرها، به دست می‌آید.

$$q_2 = \frac{\sum_{i=1}^n q_{2i}}{n}$$

$$q_{2i} = \frac{\sum_{j=1}^s wt_{ij} z_{ij}}{\sum_{j=1}^s wt_j}$$

z_{ij} امتیاز کسب شده متغیر i ام در شاخص j ام

wt_{ij} وزن شاخص j ام برای متغیر i ام

s تعداد شاخص‌هایی که برای متغیر i ام محاسبه شده است.

n تعداد متغیرهای ثبت اداری که ارزیابی شده است.

۲-۳-۱- شاخص‌های ارزیابی «متغیرهای سامانه»

شاخص‌های ارزیابی «متغیرهای سامانه» در جدول ۳ ارائه شده است.

جدول ۳- ارزیابی «متغیرهای سامانه»

ردیف	عنوان شاخص	راهنمای ارزیابی	وزن شاخص (w_e)	امتیاز کسب شده (y_1): عنوان متغیر	امتیاز کسب شده (y_2): عنوان متغیر	امتیاز کسب شده (y_3): عنوان متغیر	امتیاز کسب شده (y_n): عنوان متغیر
۱	آیا تعریف متغیر در منبع اداری با تعریف مورد نظر مرکز آمار سازگار است؟	- سازگار است (۱ امتیاز) - سازگار نیست (۰ امتیاز)	—				
۲	آیا زمان مرجع ثبت اداری قابلیت تطبیق با زمان مرجع مورد نیاز اهداف آماری را دارد؟	- کاملاً مطابقت دارد (۱ امتیاز) - با نیاز مرکز آمار، مطابقت زیادی دارد (۰,۷ امتیاز) - با نیاز مرکز آمار، مطابقت کمی دارد (۰,۳ امتیاز) - مطابقت ندارد (۰ امتیاز)	—				
۳	آیا قابلیت ارائه تفکیک‌های آماری مورد نیاز وجود دارد؟	- کاملاً مطابقت دارد (۱ امتیاز) - با نیاز مرکز آمار، مطابقت زیادی دارد (۰,۷ امتیاز) - با نیاز مرکز آمار، مطابقت کمی دارد (۰,۳ امتیاز) - مطابقت ندارد (۰ امتیاز)	—				

ردیف	عنوان شاخص	راهنمای ارزیابی	وزن شاخص (we)	امتیاز کسب شده (Y_1): عنوان متغیر	امتیاز کسب شده (Y_2): عنوان متغیر	امتیاز کسب شده (Y_3): عنوان متغیر	امتیاز کسب شده (Y_n): عنوان متغیر
۴	آیا طبقه‌بندی متغیر در منبع اداری با طبقه‌بندی مورد نظر مرکز آمار سازگار است؟	- کاملاً سازگار است (۱ امتیاز) - با نیاز مرکز آمار، سازگاری زیادی دارد یا از طریق جداول تطبیقی یا سایر روش‌ها، می‌توان سازگاری زیادی به دست آورد (۰,۷ امتیاز) - با نیاز مرکز آمار، سازگاری کمی دارد. یا از طریق جداول تطبیقی یا سایر روش‌ها، می‌توان سازگاری کمی به دست آورد (۰,۳ امتیاز) - سازگار نیست (۰ امتیاز)	—				
۵	نرخ طبقه‌بندی درست	نرخ طبقه‌بندی درست	—				
۶	نرخ پاسخی متغیر	۱- بی‌پاسخی متغیر	—				
۷	نرخ مقادیری که اشتباه یا خارج از محدوده نیستند	۱- نرخ مقادیر اشتباه یا خارج از محدوده	—				
۸	نرخ مقادیر متغیر در منبع مورد نظر با منبع دیگر	نرخ سازگاری مقادیر متغیر	—				
۹	نرخ پوشش واحد آماری (فرد، مکان، کارگاه، ...)	نرخ پوشش واحد آماری	—				

۲-۳-۲- راهنمای ارزیابی «متغیرهای سامانه»

- شاخص ۱- آیا تعریف متغیر در منبع اداری با تعریف مورد نظر مرکز آمار سازگار است؟
در صورتی می‌توان از متغیر مورد نظر استفاده کرد که تعریف ارائه شده برای آن متغیر، با تعریف مورد نظر مرکز آمار سازگار باشد. با توجه به میزان سازگاری تعاریف، ارزیابی انجام می‌شود.
تذکر: چنانچه در جدول ۱- شاخص ۱، تعریف به کار رفته مطابق با تعاریف استاندارد ارزیابی شده است، بنا بر این، سازگاری کامل وجود دارد و نمره ارزیابی ۱ درج می‌شود.

- شاخص ۲- آیا زمان مرجع ثبت اداری قابلیت تطبیق با زمان مرجع مورد نیاز اهداف آماری را دارد؟
زمان مرجع، زمانی است که داده‌ها به آن مربوط است. با توجه به میزان انطباق زمان مرجع، ارزیابی انجام می‌شود.

• شاخص ۳- آیا قابلیت ارائه تفکیک‌های آماری مورد نیاز وجود دارد؟

در صورتی که متغیر مورد نظر در منبع ثبتی وجود دارد، باید بررسی شود که آیا تفکیک‌های مورد نیاز مرکز آمار قابل ارائه است. به عنوان مثال اگر اطلاع «وضعیت مکان» در منبع ثبت وجود دارد آیا به تفکیک «محل انجام فعالیت اقتصادی»، «اقامتگاه دوم»، «خالی مسکونی»، «خالی غیرمسکونی»، «در دست ساخت یا تخریب» و «سایر» وجود دارد یا به این تفکیک قابل ارائه است. مثال دیگر می‌توان متغیر «نوع محل سکونت خانوار» را در نظر گرفت. اگر این متغیر وجود دارد، باید بررسی شود که آیا تفکیک‌های آن به صورت «واحد مسکونی آپارتمانی»، «واحد مسکونی غیر آپارتمانی»، «چادر، کپر، آلونک، زاغه یا موارد مشابه» و «سایر» است.

تذکر:

✓ تفکیک‌های جغرافیایی در این شاخص قرار نمی‌گیرد. این موضوع در شاخص ۷، بخش ارزیابی «مشخصات فنی ثبت اداری» پرداخته می‌شود.

✓ ارائه اطلاعات بر اساس طبقه‌بندی‌های بین‌الملل در شاخص ۳ بررسی می‌شود.

• شاخص ۴- آیا طبقه‌بندی متغیر در منبع اداری با طبقه‌بندی مورد نظر مرکز آمار سازگار است؟

منظور از طبقه‌بندی، طبقه‌بندی‌های بین‌المللی نظیر ISIC، ISCO، ISCED و ... است. به عنوان نمونه در ثبت‌های مرتبط با آموزش، در صورت استفاده از طبقه‌بندی، آیا از طبقه‌بندی ISCED استفاده می‌شود. در صورتی که کدگذاری در فایل انجام نشده و تنها شرح موضوع (مانند شرح رشته و مقطع تحصیلی) وجود دارد، آیا شرح موضوعات به گونه‌ای است که امکان کدگذاری بر اساس طبقه‌بندی ISCED وجود داشته باشد.

تذکر: چنانچه در جدول ۱- شاخص ۵، طبقه‌بندی به کار رفته مطابق با طبقه‌بندی استاندارد ارزیابی شده است، بنا بر این، سازگاری کامل وجود دارد و نمره ارزیابی ۱ درج می‌شود.

• شاخص ۵- نرخ طبقه‌بندی درست

در صورتی که کدگذاری توسط دستگاه اجرایی انجام شده و متن کدگذاری شده نیز موجود است، برای محاسبه این نرخ نمونه‌ای تصادفی انتخاب و سپس توسط کدگذاران ماهر، مجدد کدگذاری شده و بر اساس میزان انطباق‌ها، نرخ مورد نظر برآورد شود.

تعداد رکوردهایی که کد تخصیص داده شده به آن
توسط دستگاه اجرایی و کدگذار ماهر مرکز آمار مطابقت دارد

نرخ طبقه‌بندی درست =

تعداد کل رکوردهای بررسی شده

در صورتی که تنها متن توصیفی برای متغیر قابل کدگذاری وجود داشته باشد، نرخ طبقه‌بندی درست با محاسبه دو مقدار C_1 و C_2 به دست می‌آید.

۱. C_1 : ابتدا متن توصیفی به صورت خودکار کدگذاری می‌شود. تعداد رکوردهای دارای متن توصیفی که به صورت خودکار کدگذاری نشده‌اند یا با احتمال پایین (توسط کارشناس کدگذاری تعیین می‌شود) کدگذاری شده‌اند، C_1 نامیده می‌شود.

۲. C_2 : از بین موارد کدگذاری شده به صورت خودکار، نمونه‌ای تصادفی انتخاب و سپس توسط کدگذاران ماهر، مجدد کدگذاری می‌شود. نرخ عدم انطباق‌ها W نام‌گذاری می‌شود. مقدار C_2 از رابطه زیر محاسبه می‌شود:

$$C_2 = W(N - C_1)$$

که در آن، N تعداد رکوردهای دارای شرح کد است.

$$۳. \text{ نرخ طبقه‌بندی درست} = \frac{C_1 + W(N - C_1)}{N} - ۱$$

• شاخص ۶- نرخ پاسخی متغیر

$$۱- \text{ نرخ بی پاسخی متغیر} = \frac{\text{تعداد رکوردهای واجد شرایط پاسخگویی به متغیر مورد نظر که بی پاسخی/بی اطلاع هستند}}{\text{تعداد کل رکوردهای واجد شرایط پاسخگویی به متغیر مورد نظر}}$$

• شاخص ۷- نرخ مقادیری که اشتباه یا خارج از محدوده نیستند

تعداد رکوردهای واجد شرایط پاسخگویی به متغیر مورد نظر
که مقادیر آن متغیر، اشتباه یا خارج از محدوده است

$$۱- \text{ نرخ مقادیر اشتباه یا خارج از محدوده} = ۱ - \frac{\text{تعداد کل رکوردهای واجد شرایط پاسخگویی به متغیر مورد نظر}}{\text{تعداد کل رکوردهای واجد شرایط پاسخگویی به متغیر مورد نظر}}$$

• شاخص ۸- نرخ سازگاری مقادیر

این شاخص در دو حالت مورد محاسبه قرار می‌گیرد:

۱. نرخ سازگاری مقادیر متغیر مورد نظر با متغیر/متغیرهای دیگر در همان منبع

تعداد رکوردهایی که برای آن‌ها مقدار متغیر مورد نظر با متغیر/متغیرهای دیگر
در همان منبع دیگر سازگار است

$$\text{نرخ سازگاری مقادیر متغیر} = \frac{\text{تعداد رکوردهای منبع مورد نظر}}{\text{تعداد رکوردهای منبع مورد نظر}}$$

۲. نرخ سازگاری متغیر در منبع موردنظر با منبع دیگر

تعداد رکوردهایی که برای آن‌ها مقدار متغیر موردنظر در منبع مورد بررسی
با منبع دیگر سازگار است

$$\text{نرخ سازگاری مقادیر متغیر} = \frac{\text{تعداد رکوردهای منبع مورد نظر}}{\text{تعداد رکوردهای منبع مورد نظر}}$$

در حالت دوم:

✓ شاخص با استفاده از واحدهای مشترک دو منبع (که به وسیله فیلد یکتاکننده قابل تشخیص هستند) محاسبه می‌شود.

✓ در صورتی که مقدار متغیر برای یک واحد در منبع مورد بررسی بی‌پاسخ باشد، این واحد جزء موارد ناسازگار محسوب می‌شود.

• شاخص ۹- نرخ پوشش واحد آماری (فرد، مکان، کارگاه، ...)

با محاسبه این نرخ، میزان بیش پوششی و کم پوششی هر منبع نسبت به سایر منابع (در جامعه هدف مشترک دو فایل) مشخص می‌شود. با توجه به نوع منبع، ممکن است حداکثر دو نرخ (بیش پوششی و کم پوششی) برای یک منبع نسبت به منبع دیگر قابل محاسبه باشد. به عنوان مثال اگر A، فایل دریافتی از وزارت آموزش و پرورش و B، فایل دریافتی از ثبت احوال باشد، برای فایل A تنها نرخ بیش پوششی قابل محاسبه است؛ چرا که ممکن است فردی با وجود اینکه در جامعه هدف آموزش و پرورش قرار دارد، به دلیل ترک تحصیل یا بازماندن از آموزش، در فایل A وجود نداشته باشد. این در حالی است که برای منبع B، هم نرخ بیش پوششی و هم نرخ کم پوششی قابل محاسبه است. در صورتی که برای یک متغیر در منبع مورد بررسی، بیش از یک منبع مشابه برای مقایسه وجود داشته باشد، نرخ کم پوششی/بیش پوششی نهایی با میانگین گرفتن از نرخ‌های کم پوششی/بیش پوششی حاصل از مقایسه با منابع مختلف به دست می‌آید.

تعداد رکوردهای غیر تکراری که در منبع B وجود دارد
ولی در منبع A وجود ندارد

$$\text{نرخ کم پوششی منبع A نسبت به منبع B} = \frac{\text{تعداد رکوردهای غیر تکراری منبع A}}{\text{تعداد رکوردهای غیر تکراری منبع B}}$$

تعداد رکوردهای غیر تکراری که در منبع A وجود دارد
ولی در منبع B وجود ندارد

$$\text{نرخ بیش پوششی منبع A نسبت به منبع B} = \frac{\text{تعداد رکوردهای غیر تکراری منبع A}}{\text{تعداد رکوردهای غیر تکراری منبع B}}$$

این شاخص دوبار محاسبه می‌شود:

۱. پس از دریافت اولیه فایل از منبع مورد بررسی

۲. پس از بررسی تناقض‌ها و رفع مشکلات

در مثال قبل، این نرخ یکبار پس از شناسایی رکوردهای غیر تکراری که در منبع آموزش و پرورش وجود دارد ولی در منبع ثبت احوال وجود ندارد، محاسبه می‌شود. در مرحله دوم، با ارسال فایل تناقض‌ها به ثبت احوال، امکان دارد تعدادی از رکوردها، تغییر وضعیت داشته باشند، به عنوان مثال، فردی فوت شده و به این دلیل در فایل ثبت احوال وجود ندارد. پس از بررسی این موارد، نرخ، مجدد محاسبه می‌شود.

۲-۴- ارزیابی میزان قابل استفاده بودن رکوردهای ثبت اداری (q_3)

شاخص q_3 «نرخ قابل استفاده بودن رکوردهای ثبت اداری» است. این شاخص پس از دریافت فایل داده محاسبه می‌شود. مقدار شاخص q_3 از رابطه زیر به دست می‌آید.

$$q_3 = \frac{a-(b+c)}{a}$$

که در آن:

a: تعداد رکوردهای موجود در فایل

b: تعداد رکوردهای تکراری (تمام فایل‌ها، چه کد شناسایی یکتا داشته باشند چه نداشته باشند)

c: تعداد رکوردهای بدون کد شناسایی یکتا (مقدار c برای فایل‌هایی که کد شناسایی یکتا ندارند، صفر در نظر گرفته می‌شود)

✓ برای محاسبه b، فرض شود که رکوردی ۴ بار در فایل وجود دارد، در این صورت مقدار b برابر ۴ است.

✓ برای محاسبه شاخص q_3 ابتدا باید تعداد رکوردهای مشترک b و c را حذف کرده و سپس شاخص را محاسبه نمود.

۲-۵- محاسبه کیفیت ثبت اداری و میزان قابل استفاده بودن آن ($q_{\alpha 1}$ و $q_{\alpha 2}$)

- $q_{\alpha 1}$ مؤلفه «کیفیت ثبت اداری» است که بر اساس شاخص‌های معرفی شده در بخش‌های قبل و جدول ۴ محاسبه می‌شود.
- $q_{\alpha 2}$ مؤلفه «قابلیت تولید ثبت آماری» است که بر اساس شاخص‌های معرفی شده در بخش‌های قبل و جدول ۴ محاسبه می‌شود.

دو مؤلفه $q_{\alpha 1}$ و $q_{\alpha 2}$ با توجه به شاخص‌های معرفی شده در جداول ۱ تا ۳ و نقش شاخص‌ها برای هر مؤلفه طبق جدول ۴ محاسبه می‌شود.

$$q_{\alpha i} = \frac{\sum_{i=1}^m w_i x_i + \sum_{i=1}^r w e_i y_i + \sum_{i=1}^n \sum_{j=1}^s w t_{ij} z_{ij} + q_3}{\sum_{i=1}^m w_i + \sum_{i=1}^r w e_i + \sum_{i=1}^n \sum_{j=1}^s w t_{ij} + 1}$$

x_i امتیاز کسب شده ثبت اداری در شاخص آم (در جدول ۱)

w_i وزن شاخص آم (در جدول ۱)

m تعداد شاخص‌هایی که برای ثبت اداری مورد نظر محاسبه شده است (در جدول ۱)

y_i امتیاز کسب شده ثبت اداری در شاخص آم (در جدول ۲)

$w e_i$ وزن شاخص آم (در جدول ۲)

r تعداد شاخص‌هایی که برای ثبت اداری مورد نظر محاسبه شده است (در جدول ۲)

z_{ij} امتیاز کسب شده متغیر آم در شاخص زام (در جدول ۳)

$w t_{ij}$ وزن شاخص زام برای متغیر آم (در جدول ۳)

s تعداد شاخص‌هایی که برای متغیر آم محاسبه شده است (در جدول ۳)

n تعداد متغیرهای ثبت اداری که ارزیابی شده است (در جدول ۳)

q_3 : نرخ قابل استفاده بودن رکوردهای ثبت اداری

شاخص‌های ارزیابی در یک نگاه

جدول ۴ شاخص‌های ارزیابی را به تفکیک نوع ارزیابی نشان می‌دهد.

ردیف	شاخص ارزیابی	ارزیابی کیفیت ثبت اداری	ارزیابی قابلیت تولید ثبت آماری
جدول ۱- ارزیابی کیفیت ثبت داده‌ها (q.)			
۱	آیا عنوان متغیر، تمامی رخداد‌های مربوطه را پوشش می‌دهد؟	✓	
۲	آیا تعاریف و مفاهیم، کامل و مقایسه‌پذیر است؟	✓	
۳	آیا استانداردها و قرارداد‌های مشترک در مورد زمان (از قبیل سری زمانی، بازه زمانی و مقطع زمانی) برای ثبت متغیرهای مورد نظر رعایت می‌شود؟	✓	
۴	آیا استانداردها و قرارداد‌های مشترک در مورد مشخصات مکانی (از قبیل سطوح جغرافیایی، یکپارچگی داده‌ها، ساختار و قابلیت استفاده در سیستم‌های اطلاعات جغرافیایی GIS و...) برای ثبت متغیرهای مورد نظر رعایت می‌شود؟	✓	
۵	آیا استانداردها و قرارداد‌های مشترک در مورد تفکیک‌ها (از قبیل سن، مدرک تحصیلی و...) و طبقه‌بندی‌های بین‌المللی (از قبیل ISCED, ISIC و...) برای ثبت قلم آماری مذکور رعایت می‌شود؟	✓	
۶	آیا داده‌ها در پایین‌ترین نقطه، ثبت می‌شوند؟	✓	
۷	آیا داده‌ها به هنگام وقوع عملیات ثبت می‌شوند؟	✓	
۸	آیا ورود اطلاعات به سیستم همزمان با انجام عملیات ثبت صورت می‌گیرد؟	✓	
۹	آیا عمل ثبت در فرآیند تولید توسط افراد مسئول و پاسخگو انجام می‌شود؟	✓	
۱۰	آیا فرم‌های عملیاتی که در فرآیند تولید اقلام آماری به کار می‌روند برای تأمین اطلاعات مورد نیاز طراحی مناسبی دارند؟	✓	
۱۱	آیا برای کنترل و تأیید صحت ثبت در مراحل مختلف، مکانیزمی وجود دارد؟	✓	
۱۲	آیا الزام قانونی برای ثبت متغیر وجود دارد؟	✓	
۱۳	آیا عمل انتقال در فرآیند تولید توسط افراد مسئول و پاسخگو انجام می‌شود؟	✓	
۱۴	آیا داده‌های ثبتی تولید شده از مسیرهای مناسب به سطوح بالاتر منتقل می‌شوند؟	✓	
۱۵	آیا در مسیر انتقال، هر ایستگاه وظایف خود (از قبیل کنترل، تجمیع درست یا اقدام به موقع برای انتقال اطلاعات و...) را به خوبی انجام می‌دهد؟	✓	
۱۶	آیا تکنولوژی انتقال متناسب با قلم آماری و اهمیت و فرآیند تولید آن می‌باشد؟	✓	
۱۷	آیا امنیت داده‌ها و اطلاعات هنگام انتقال رعایت می‌شود؟	✓	
جدول ۲- ارزیابی کیفیت فایل داده‌های دریافتی (q_۱)			
۱	آیا فایل‌های داده ارسالی با موضوع یکسان، یکپارچه است؟		✓
۲	آیا مدل داده برای فایل داده در اختیار مرکز قرار گرفته است؟		✓
۳	آیا فراداده‌های مورد نیاز، همراه با داده‌ها در اختیار مرکز قرار گرفته است؟		✓
۴	آیا موارد اصلاحی مورد نظر مرکز (نظیر ساختار، محتوا و نوع فایل)، در دوره‌های بعدی اعمال شده است؟		✓

ردیف	شاخص ارزیابی	ارزیابی کیفیت ثبت اداری	ارزیابی قابلیت تولید ثبت آماری
۵	آیا ساختار و نوع فایل، در دوره‌های مختلف ارسال یکسان است؟		✓
۶	آیا هر موجودیت در فایل داده ارسالی، دارای متغیر/ متغیرهای یکتاکننده است؟	✓	✓
۷	آیا فایل داده دارای متغیر/ متغیرهای اطلاعاتی با قابلیت پیوند با فایل‌های داده سایر دستگاه‌های اجرایی است؟	✓	✓
۸	آیا ثبت اداری قابلیت پیوند با اطلاعات مکانی برای تفکیک‌های جغرافیایی را دارا است؟		✓
۹	آیا داده‌ها بر اساس زمان‌بندی مورد توافق، دریافت شده است؟		✓
جدول ۳- ارزیابی متغیرهای سامانه (qr)			
۱	آیا تعریف متغیر در منبع اداری با تعریف مورد نظر مرکز آمار سازگار است؟		✓
۲	آیا زمان مرجع ثبت اداری قابلیت تطبیق با زمان مرجع مورد نیاز اهداف آماری را دارد؟		✓
۳	آیا قابلیت ارائه تفکیک‌های آماری مورد نیاز وجود دارد؟		✓
۴	آیا طبقه‌بندی متغیر در منبع اداری با طبقه‌بندی مورد نظر مرکز آمار سازگار است؟		✓
۵	نرخ طبقه‌بندی درست	✓	✓
۶	نرخ پاسخی متغیر	✓	✓
۷	نرخ مقادیری که اشتباه یا خارج از محدوده نیستند	✓	✓
۸	نرخ سازگاری مقادیر متغیر در منبع موردنظر با منبع دیگر	✓	✓
۹	نرخ پوشش واحد آماری (فرد، مکان، کارگاه، ...)	✓	✓
ارزیابی میزان قابل استفاده بودن رکوردهای ثبت اداری (qr)			
۱	نرخ قابل استفاده بودن رکوردهای ثبت اداری	✓	✓

منابع

- [۱] قادری، سمانه؛ حق گو، نصرت‌ا.. و حقیری، مریم (۱۳۹۶)، بررسی مطالعات انجام شده در زمینه ارزیابی کیفیت آمارهای ثبتی، دفتر استانداردها و نظارت بر طرح‌های آماری.
- [۲] سید نعمت اله میرفلاح نصیری، طرح پژوهشی «ارزیابی کیفیت داده‌های ثبتی یا اداری برای استفاده در سرشماری ثبتی مبنا»
- [۳] طالبی، مریم؛ طرح پژوهشی «طراحی ابزار و راهکار اجرایی ارزیابی و اندازه‌گیری کیفیت آمارهای ثبتی»؛ بهار ۱۳۹۲.
- [۴] مرکز آمار ایران، ارزیابی کیفیت اقلام آمار ثبتی دستگاه‌های اجرایی کشور
- [۵] مرکز آمار ایران، دستورالعمل ارزیابی کیفیت اقلام آمار ثبتی دستگاه‌های اجرایی کشور
- [۶] Asamer, Astleithner, Cetkovi, Humer, Lenk, Moser, Rechta (۲۰۱۶), Quality Assessment for Register-based Statistics -Results for the Austrian Census ۲۰۱۱, Austrian Journal of Statistics, March ۲۰۱۶, Volume ۴۵, ۳-۱۴.
- [۷] BLUE-Enterprise and Trade Statistics, (BLUE-ETS), Deliverable ۴.۱. List of quality groups and indicators identified for administrative data sources. ۱۰ March ۲۰۱۱.
- [۸] Dass. Piet, Ossen. Saskia, Vis-Visschers. Rachel, Arends-Toth. Judit (۲۰۰۹), Checklist for the Quality evaluation of Administrative Data Sources, Statistics Netherlands, ISSN: ۱۵۷۲-۰۳۱۴The
- [۹] Lavigne Mylène and Nadeau Christian (۲۰۱۴), A Framework for the Evaluation of Administrative Data, Proceedings of Statistics Canada Symposium ۲۰۱۴, Beyond Traditional Survey Taking: Adapting to a Changing World
- [۱۰] Reinert Regin (۲۰۱۶), Evaluating the Quality of Administrative Data as Input for Official Statistics,
- [۱۱] Statistics Denmark, European Conference on Quality in Official Statistics (Q۲۰۱۶) Madrid, ۳۱ May-۳ June ۲۰۱۶ ۲۰۰۹
- [۱۲] Statistics Netherlands (۱۹۹۸), A model for evaluating the quality of business surveys.