

Causal Mapping in Managing Statistical Outputs with Data Quality Approach

Masood Ghafari

PhD Candidate; Faculty of Operational Management;
Allameh Tabataba'i University; Tehran, Iran m.ghafari@atu.ac.ir

Adel Azar

PhD in Operation Research; Professor; Tarbiat Modarres University;
Tehran, Iran;
Corresponding Author azara@modares.ac.ir

Ashkan Shabbak

PhD in Applied Statistics; Assistant Professor; Statistical Research
and Training Center; Tehran, Iran ashkan@srtc.ac.ir

Received: 08, Apr. 2017 Accepted: 24, Jun. 2017

Abstract: Data quality is an important criterion for each database. Data without quality cause less efficiency and effectiveness in decision-making. Organizations will face a harmful crisis in this situation and condition. However, quality assurance is a new approach in quality management. This approach continues to data quality management and quality assurance is first until end of collection data. The aim of this paper is recognition dimensions of data quality and their concepts. Managing statistical outputs is one of the important dimensions of data quality that has seven elements. This research wants to explore the impact of them. The method of this paper is DEMATEI. Soft operational research have many methods and one of them is DEMATEI. The results of this article show that assuring accuracy and completeness have power of influence. They are important elements that managers should take notice of.

Keywords: Data Quality, Data Quality Management, Managing Statistical Outputs, Causal Mapping, DEMATEI

Iranian Journal of
Information
Processing and
Management

Iranian Research Institute
for Information Science and Technology
(IranDoc)

ISSN 2251-8223

eISSN 2251-8231

Indexed by SCOPUS, ISC, & LISTA

Vol. 33 | No. 3 | pp. 1041-1064

Spring 2018

<https://doi.org/10.35050/JIPM010.2018.040>



نگاشت علی مدیریت محصول آماری با رویکرد کیفیت داده

مسعود غفاری

دانشجوی دکتری تحقیق در عملیات؛ مدیریت صنعتی؛
دانشگاه علامه طباطبائی | m.ghafari@atu.ac.ir

عادل آذر

دکتری تحقیق در عملیات؛ استاد؛
دانشگاه تربیت مدرس؛
پدیدآور رابط | azara@modares.ac.ir

اشکان شباک

دکتری آمار کاربردی؛ استادیار؛
پژوهشکده مرکز آمار ایران | ashkan@src.ac.ir



دریافت: ۱۳۹۶/۱/۱۹ | پذیرش: ۱۳۹۶/۰۴/۰۳ | مقاله برای اصلاح به مدت ۱۴ روز نزد پدیدآوران بوده است.

فصلنامه | علمی پژوهشی
پژوهشگاه علوم و فناوری اطلاعات ایران
(ایرانداک)

شاپا (جایی) ۲۲۵۱-۸۲۲۳

شاپا (الکترونیکی) ۲۲۵۱-۸۲۳۱

نمایه در SCOPUS، ISC، LISTA، و

jipm.irandoc.ac.ir

دوره ۳۳ | شماره ۳ | صص ۱۰۶۱-۱۰۶۴

بهار ۱۳۹۷

<https://doi.org/10.35050/JIPM010.2018.040>



چکیده: کیفیت داده یک ویژگی مهم و ضروری برای هر پایگاه داده است. کیفیت داده از جمله مباحث مهمی است که اگر به آن توجه نشود، نمی‌تواند نیاز سازمان را تأمین کند. یکی از جنبه‌های مهم در مورد کیفیت داده، تضمین کیفیت است. شاید داده‌های کافی جمع‌آوری شده باشند، اما این که داده‌ها از دقت کافی برخوردار باشند و در عین حال، به‌روز بوده و سازمان‌ها، از جمله سازمان‌های ملی آماری، بتوانند برای تصمیم‌گیری و تصمیم‌سازی از آن استفاده کنند، مربوط به جنبه تضمین کیفیت داده‌هاست. هدف پژوهش حاضر آن است که ضمن بازشناسی مفهوم مدیریت کیفیت داده و ابعاد آن، به بُعد مدیریت محصول آماری در مدیریت کیفیت داده پردازد. در این پژوهش ضمن شناسایی عناصر بُعد مدیریت محصول آماری، به تبیین شدت اثر هر یک از عناصر پرداخته شده است. روش تحقیق این پژوهش، «دیمتل» است که از جمله روش‌های حوزه تحقیق در عملیات نرم می‌باشد. یافته‌ها حاکی از آن است که عناصر تضمین انسجام، تضمین درستی و جامعیت از شدت اثر بالایی برخوردار بوده و نسبت به سایر عناصر بُعد مدیریت محصول آماری از اثرگذاری بالایی برخوردار است. لذا، مدیران مراکز آماری جهت افزایش و ارتقاء سطح کیفیت داده‌ها لازم است عناصر مذکور را در اولویت برنامه‌ریزی خود قرار دهند.

کلیدواژه‌ها: کیفیت داده، مدیریت کیفیت داده، مدیریت محصول آماری، نگاشت علی، دیمتل

۱. مقدمه

اهمیت مسئله کیفیت داده موضوعی بدیهی و آشکار است، اما اگر یک‌سری داده در مقابل کسی گذاشته شود، وی نمی‌تواند در مورد باکیفیت بودن یا صحت و درستی آن قضاوت درستی داشته باشد (Herzog, Scheuren & Winkler 2007). هنگامی که از مردم درباره کیفیت داده پرسش می‌شود، بیشتر افراد تنها به دقت آن توجه می‌کنند، در حالی که اگر از میان ابعاد کیفیت داده تنها بعد دقت آن مد نظر قرار گیرد، این نگاه بسیار ساده‌اندیشی خواهد بود. به عبارت دیگر، درباره کیفیت داده ابعاد دیگری مطرح است که دقت تنها یکی از آنهاست (Batini & Scannapieca 2006, 4-5).

در حال حاضر، یکی از مهم‌ترین مشکلات مسئولان کشور برای برنامه‌ریزی و تصمیم‌گیری اقتصادی، اجتماعی و فرهنگی کمبود داده و ارقام گویا و قابل اعتماد است. این معضل از دیرباز و از شروع برنامه‌ریزی در ایران به‌عنوان یک عامل بازدارنده وجود داشته است و اکنون با گذشت حدود هشت دهه، هنوز از آن به‌عنوان یکی از مهم‌ترین مشکلات برنامه‌ریزی در کشور یاد می‌شود. یکی از دلایل کمبود داده در این زمینه، ساختار نارسای نظام آماری کشور است که موجب شده آمار و ارقام دقیق، بی‌طرف، و به‌هنگام کمتر در اختیار برنامه‌ریزان و محققان قرار گیرد و آمار و ارقام موجود در کشور تکافوی نیاز مسئولان و برنامه‌ریزان را نکند (گزارش دفتر بررسی‌های اقتصادی ۱۳۷۵، ۳۴). به‌طور کلی، چنانچه داده‌های مورد نظر از سطح کیفیتی پایینی برخوردار باشند، سازمان‌ها را در سه سطح عملیاتی، تکنیکی و استراتژیکی با مشکل مواجه خواهند کرد (Pollock 2012, 9).

علی‌رغم هزینه و تلاش‌های زیادی که برای جمع‌آوری و تجزیه و تحلیل داده صرف می‌شود، در حال حاضر، داده‌های بسیاری از سازمان‌ها از سطح استاندارد کافی جهت تصمیم‌گیری برخوردار نیستند. هرچند در جهان امروز نرم‌افزارهای قدرتمند آماری وجود داشته و امکان تحلیل‌های آماری با ضریب اطمینان بالایی وجود دارند، اما این امر با وجود داده‌های آشفته و غیردقیق، فایده چندانی ندارد (Herzog, Scheuren & Winkler 2007, 8).

این در حالی است که داده‌ها در هر سازمان حکم آب را دارند و برای بقای هر سازمانی ضروری هستند. از سوی دیگر، همان‌گونه که بررسی کیفیت آب از اهمیت

دوچندانسی برخوردار است و چنانچه آب مورد نظر استانداردهای لازم را نداشته باشد می‌تواند جان چندین هزار انسان را بگیرد، سازمان نیز در صورت عدم برخورداری از داده‌ها و اطلاعات با کیفیت قادر به ادامه حیات نخواهد بود (Morbey 2013, 3). کیفیت داده یک ویژگی مهم و ضروری برای هر پایگاه داده است؛ چرا که تضمین سلامت و صحت داده‌ها هنگام ترکیب و یکپارچگی آن‌ها، از لازمه‌های اساسی مدیریت کیفیت داده است. پژوهشگران حوزه کیفیت در تلاش‌اند موضوع و اهداف مورد نظر در اندازه‌گیری کیفیت داده را توسعه ببخشند (Pollock 2012, 6).

داده‌ها زمانی از کیفیت لازم برخوردار خواهند بود که متناسب با مسئله جمع‌آوری و پردازش شوند. این امر توسط تصمیم‌گیرندگان مراکز آماری و دیگر مجریان مشخص خواهد شد. در تولیدات روزانه کارخانه‌ها، محصول، زمانی با کیفیت تلقی خواهد شد که با استانداردهای از پیش تعیین‌شده متناسب باشد. پس، هنگامی که محصولات مطابق با استانداردها باشند، موفقیت حاصل شده است (Herzog, Scheuren & Winkler 2007, 7-8). بنا بر دلایل فوق، مراکز ملی آمار در موقعیتی قرار گرفته‌اند که لازم است درباره کیفیت داده‌های آماری تولیدی خود تلاش جدی نمایند. در واقع، از اواسط دهه ۱۹۸۰، اداره‌های ملی و سازمان‌های بین‌المللی آمار با توجه بیشتری به کیفیت آماری و بهبود آن پرداخته‌اند (دن میونگ لی، ۱۳۸۳، ص. ۶۹).

۲. پیشینه نظری

۲-۱. کیفیت داده

اولین گام برای بهبود کیفیت داده‌ها و ارتقای آن‌ها توجه به معنای کیفیت داده است. سازمان در درجه اول باید بداند که کیفیت داده چیست تا در مراحل بعدی به فکر توسعه آن باشد (Wang & Strong 1996, 6).

مفهوم کیفیت به‌طور کلی، امری پیچیده و در عین حال چندبعدی است. این مفهوم زمانی که محصول، داده‌های آماری باشد، چه بسا پیچیده‌تر هم می‌شود. گوناگونی و گستره مشتریان سازمان‌های آماری و کاربران محصولات آماری از سویی و ویژگی‌های منحصر به فرد آمار از دید محصول از سوی دیگر، به این پیچیدگی افزوده است و این امر موجب دشواری کیفیت برای این محصول می‌شود. از این رو، در سال‌های گذشته برای

تعریف و استانداردسازی مفهوم کیفیت در داده‌های آماری و مدل‌سازی مدیریت کیفیت در سازمان‌هایی که به تولید و انتشار آمار و اطلاعات می‌پردازند، تلاش‌های فراوانی به عمل آمده است (شباک ۱۳۹۲، ۵).

به‌طور کلی، در تعریفی که از کیفیت داده ارائه شده، آن را متناسب با نیاز مورد نظر کاربر تعریف کرده‌اند. در این تعریف، مفهوم کیفیت داده باید با نیاز کاربران متناسب باشد (Tayi & Ballou 1998, 54). هر قدر داده‌های مورد نظر بتواند اهداف بیشتری از نیاز کاربران را محقق نماید، از سطح کیفیتی بالاتری برخوردار هستند. آنچه در این تعریف قابل بررسی است آن است که کیفیت داده ارتباط نزدیکی با کاربر و نیاز او دارد (Larburu et al. 2015, 452). ابعاد گسترده‌ای از کیفیت داده‌ها مطرح شده؛ مانند: دقت، اطمینان، به‌هنگامی، تناسب، جامعیت، سازگاری و ... (Haug et al. 2013, 236) هر بُعد از ابعاد کیفیت داده دربرگیرنده ویژگی‌های مشخصی است که در نهایت، مفهوم کیفیت داده را شکل می‌دهند (Batini & Scannapieca 2006, 19) که در ادامه به تفصیل به آن‌ها پرداخته می‌شود. از منظر محققان، موضوع کیفیت داده حوزه‌های مختلفی را شامل می‌شود که می‌توان از میان آن‌ها به مواردی همچون علم مدیریت، علم آمار و علم فناوری اطلاعات اشاره نمود. تئوری‌ها نشان می‌دهند که ظهور نظریه‌های کیفیت داده به سال ۱۹۶۰ باز می‌گردد. این نظریه‌ها در مدیریت تداوم داشته و از سال ۱۹۸۰ با مطرح شدن مسائل مربوط به کنترل داده‌های سیستم تولید جهت شناسایی کالاهای معیوب مورد توجه بیشتری قرار گرفت. در سال ۱۹۹۰، همراه با علم کامپیوتر مسائلی همچون اندازه‌گیری، ارتقا، بهبود، و تعریف کیفیت در داده‌های الکترونیکی در پایگاه‌های داده و سیستم‌های پشتیبان این حوزه قوت بیشتری گرفت (Batini & Scannapieca 2006, 4).

۲-۲. مدیریت کیفیت داده

همان‌گونه که در بسیاری از صنایع و سازمان‌ها جهت تولید محصولات با کیفیت رویکردی به نام مدیریت کیفیت جامع^۱ وجود دارد که در آن قواعد، راهنمایی‌ها و تکنیک‌های تولید یک محصول و یا خدمت با کیفیت وجود دارد، جهت ارائه اطلاعات جامع و با کیفیت بالا رویکردی به نام مدیریت کیفیت داده جامع وجود دارد. مطابق

1. Total Quality Management (TQM)

با چرخه «دینگ»^۱ که شامل مراحل برنامه‌ریزی، اجرا، کنترل و ارتقا است، در مدیریت کیفیت داده جامع نیز تعریف، اندازه‌گیری، تجزیه و تحلیل و ارتقا وجود دارد. در این چرخه، ابتدا کیفیت داده و ابعاد آن به صورت دقیق تعریف شده و در مرحله بعد، اندازه‌گیری واحدها بررسی می‌شود. در مرحله سوم، به تجزیه و تحلیل داده‌ها و عواملی که موجب تضعیف کیفیت آن‌ها شده، پرداخته می‌شود. در مرحله چهارم، به رفع نواقص و عوامل مزاحم پرداخته شده و به اعمال تکنیک‌هایی جهت ارتقای آن‌ها پرداخته می‌شود (Leonowich-Graham 2003, 2-3).

برای اجرای مدیریت کیفیت داده در هر سازمانی باید چارچوب زیر در نظر گرفته شود:

۱. کلیه مراحل تولید داده باید مشخص و قابل درک باشند؛
۲. تیم مدیریت کیفیت داده، که شامل افرادی همچون مدیر عالی سازمان، مدیر فناوری اطلاعات، اعضای جمع‌آوری اطلاعات، مشتریان و سرپرستان پایگاه‌های داده است، تشکیل شود؛
۳. ارزیابی کیفیت داده و مهارت‌های مدیریت کیفیت داده به اعضای مذکور آموزش داده شود؛
۴. تیم مدیریت کیفیت داده در سازمان تثبیت شود (Wang, Ziad, & Lee 2001, 6).

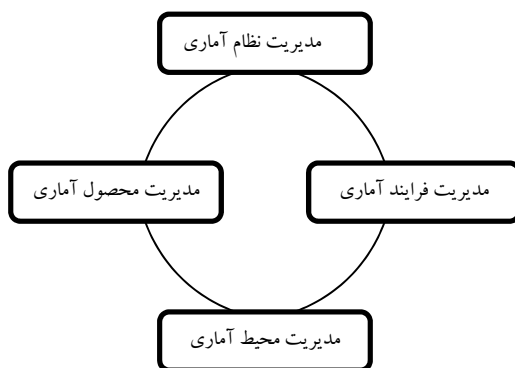
۲-۳. چارچوب ملی تضمین کیفیت

همان‌گونه که ملاحظه شد، مرحله اول در مدیریت کیفیت داده، تعریف و بررسی ابعاد کیفیت آن است. نظام‌های گوناگونی پیرامون ارائه تعریفی جامع و یکپارچه از ابعاد کیفیت داده مطرح شده است. یکی از این چارچوب‌ها، «چارچوب ملی تضمین کیفیت» آماری^۲ است. «چارچوب ملی تضمین کیفیت» آماری، چارچوبی است که یک سازمان آماری ملی را به سوی ارائه محصولی با کیفیت رهنمون ساخته و چون چتری همه فعالیت‌های وابسته به کیفیت را در سازمان دربرمی‌گیرد. از این رو، به آن چارچوب کیفیت گفته می‌شود. این مفهوم دربردارنده مؤلفه‌های تضمین، ارزیابی و مستندات کیفیت است. منظور از تضمین کیفیت، همه برنامه‌ریزی‌ها و اقدامات نظام‌مندی است که برای اطمینان از منطبق بودن

1. Deming

2. Natioan Quality Framework Assurance (NQFA)

فرایندهای کاری و رویه‌های جاری با الزامات و استانداردهای از پیش طراحی شده، انجام می‌شود تا محصول نهایی (آمار) دقیقاً مطابق با ویژگی‌های پیش‌بینی شده، تولید شود. منظور از ارزیابی کیفیت نیز ارزیابی و بازرسی‌هایی است که به‌طور کلی، کیفیت محصول آماری و داده‌ها را با توجه به معیارهای کیفی یا کمی تعیین شده کنترل می‌کند. مستندات کیفیت نیز شامل همهٔ مقررات، تعاریف و مفاهیم، استانداردها، گزارش‌های ارزیابی و دیگر مستنداتی است که در نظام مدیریت کیفیت استفاده شده و یا تولید می‌شود (NQAF Glossary 2012, 1-2).



شکل ۱. ابعاد چارچوب ملی تضمین کیفیت (4, 2012 NQAF)

تهیه و تدوین چارچوب کیفیت برای مراکز ملی آماری به اندازه‌ای اهمیت دارد که «کمیسیون آمار سازمان ملل» در چهل و یکمین نشست خود در سال ۲۰۱۰ گروهی را مؤلف به تهیه و تدوین چارچوب ملی تضمین کیفیت نمود. در طراحی الگوی NQAF، مقررات آماری «اروپا»، چارچوب ارزیابی کیفیت داده‌های «صندوق پول»، چارچوب تضمین کیفیت «کانادا»، آن‌گونه که در گزارش ۲۰۱۰ این مرکز به «کمیسیون آمار» ارائه شده، و سند موسوم به ساختارهای پیشنهادی مقررات عملکرد مطلوب آماری برای کشورهای آمریکای لاتین، مورد استفاده قرار گرفت و سعی شد که الگوی مذکور تا حد ممکن با آن‌ها مطابقت داشته باشد. این چارچوب در سال ۲۰۱۲ نهایی شد (NQAF Glossary 2012, 1-2).

۱-2. «چارچوب ملی تضمین کیفیت» (NQAF) در چهار بعد اقدام به بیان اصول و راهنمای خود کرده است که عبارت‌اند از:

۱. بعد مدیریت محصول آماری؛

۲. بعد مدیریت نظام آماری؛

۳. بعد مدیریت فرایند آماری؛

۴. بعد مدیریت محیط آماری.

هدف پژوهش حاضر در وهله اول شناسایی عناصر بُعد مدیریت محصول آماری در مدیریت کیفیت داده است. منظور از محصول آماری، نتایج حاصل از فرایندهای آماری است که در دسترس کاربران قرار می‌گیرد. محصول آماری می‌تواند به شکل آمارها، تحلیل‌های آماری و داده‌های خرد باشد (NQAF Glossary 2012, 21). مدیریت محصول آماری نقش اساسی در مدیریت کیفیت داده ایفا می‌کند؛ چرا که مبتنی بر نیاز کاربران، سازمان‌های دولتی، مؤسسات تجاری، پژوهشگران و ... است. کلیه نهادهای مذکور با بُعد مدیریت محصول آماری بیشترین تعامل را خواهند داشت (ESS Standard for Quality Reports 2009, 19). پس از شناسایی عناصر تشکیل‌دهنده بُعد مدیریت محصول آماری، به تبیین شدت اثر هر یک از آن عناصر پرداخته خواهد شد که در ادامه مقاله مورد بررسی قرار خواهد گرفت. بنابراین، سؤالات مقاله حاضر از قرار زیر است:

۱. عناصر بُعد مدیریت محصول آماری در مدیریت کیفیت داده چیست؟ چه عناصری

بُعد مدیریت محصول آماری را شکل می‌دهند؟

۲. نگاشت علی و شدت اثر هر یک از این عناصر به چه صورت است؟ کدام یک

از عناصر مدیریت محصول آماری از شدت اثر بالایی برخوردار هستند؟

با توجه به هدف پژوهش حاضر، ابتدا به معرفی و شرح عناصر مدیریت محصول

آماری که از مطالعات کتابخانه‌ای به دست آمده‌اند، پرداخته می‌شود.

۲-۴. تضمین مناسبت^۱

یکی از مهم‌ترین عناصر مدیریت محصول آماری، تناسب داده‌ها با نیاز و مسئله

سازمان‌ها و کاربران است (Herzog, Scheuren & Winkler 2007, 8). مناسبت یعنی میزان و

درجه‌ای که اطلاعات و داده‌های آماری تولیدشده، نیاز فعلی و یا بالقوه کاربران را برطرف

می‌نماید (NQAF Glossary 2012, 17-18). تضمین مناسبت شامل فرایندی است که در آن

1. relevance

مشخص می‌شود سازمان چه داده‌هایی و به کدام نوع آن نیاز دارد. نیاز سازمان در اولویت اصلی این جمع‌آوری داده قرار دارد (Statistics Canada's Quality Assurance Framework, 2002, 5).

تضمین مناسبت از آن جهت مهم است که کلیه برنامه‌های سازمان و برآورد سایر خروجی‌ها وابسته به آن است. از آنجا که نیازها در طول زمان تغییر می‌کنند، این فرایند باید به‌طور پیوسته مورد بازنگری واقع شده و با نیازهای سازمان منطبق باشد. در این انطباق محدودیت‌هایی وجود دارد؛ به‌عنوان مثال، می‌توان به وابستگی متقابل بین برنامه‌های مختلف اشاره کرد. در بسیاری از موارد، داده‌های مربوط به یک برنامه از برنامه‌ای دیگر تغذیه می‌شود. داده‌های مربوط به فروش سالیانه در یک کشور، از محل داده‌های مربوط به تولید ناخالص داخلی استفاده می‌شود. بنابراین، اثر انطباق یک برنامه بر سایر برنامه‌ها را باید در نظر گرفت (همان).

۲-۵. تضمین درستی^۱

منظور از درستی داده‌ها، درجه نزدیکی مقادیر و اندازه‌های محاسبه‌شده یا برآوردشده از یک سنج به مقدار واقعی آن سنج است. این بُعد در حقیقت خطاهای نمونه‌گیری شده و نشده را در برمی‌گیرد. درستی در کیفیت داده، انعکاسی از توصیف صحیح پدیده‌هاست (NQAF 2012, 38). در فرایند مدیریت کیفیت داده، تضمین درستی به شناسایی و کنترل خطاهای بالقوه در داده‌ها می‌پردازد. به عبارت دیگر، از طریق فرایند تضمین درستی، از ورود داده‌های مخرب جلوگیری به عمل می‌آید (Statistics Canada's Quality Assurance Framework 2002, 10). هرچه دقت و درستی بیشتری در جمع‌آوری داده‌ها لحاظ شود، بدیهی است که هزینه بیشتری بر سازمان وارد خواهد شد؛ در حالی که شاید سازمان نتواند از عهده این هزینه‌ها برآید و متناسب با بودجه اختصاصی آن نباشد. پس، کیفیت دقت داده‌ها باید متناسب با نیاز سازمان و بودجه مالی آن باشد (Herzog, Scheuren & Winkler 2007, 8-9). شاید وابسته بودن کیفیت داده‌ها به دقت و درستی، به خودی خود مسئله کوچکی باشد که مدیر عملیاتی در فرایند جمع‌آوری داده‌ها باید به آن توجه کند، اما تضمین درستی آن جایی اهمیت ویژه پیدا می‌کند که جمع‌آوری داده‌ها مستلزم

1. accuracy

صرف هزینه و انرژی بالایی بوده و چنانچه به آن توجه نشود، موجب اتلاف منابع خواهد شد. بنابراین، دقت و درستی داده‌ها اهمیت والایی در مدیریت کیفیت داده داشته و باید در مراکز آماری لحاظ شود. لازم به ذکر است که هر قدر تلاش شود، باز نمی‌توان از رخداد کلیه خطاها در داده‌ها جلوگیری به عمل آورد. در این هنگام باید به این سؤال پاسخ داده شود که متغیر اصلی و هدف در ارائه داده‌ها چیست؟ در جمع‌آوری داده‌های مورد نظر چه میزانی از دقت لازم است؟ با پاسخ به سؤالات فوق می‌توان به صورت هدفمندتر به فرایند تضمین درستی پرداخت (Statistics Canada's Quality Assurance Framework 2002, 12).

۲-۶. تضمین به‌هنگامی و به موقع بودن^۱

این عنصر از بُعد مدیریت محصول آماری بیان می‌کند که اطلاعات حاصل از داده‌ها زمانی قابل ارجاع خواهد بود که دوره داده‌ها به تأخیر نیفتاده باشد و همچنین، تاریخ و زمان داده‌ها مشخص باشد. به عبارت دیگر، عنصر به‌هنگام بودن بر این مسئله تأکید می‌کند که بین رخداد پدیده‌ها و توصیف آن‌ها نباید زمان زیادی بگذرد (ESS Standard for Quality Reports 2009, 19).

به گزارش «مرکز ملی آمار استرالیا» عنصر به‌هنگامی در موارد زیر بر پایگاه داده اثر می‌گذارد:

۱. هنگامی که کاربر به داده‌های سال جاری نیاز دارد، ولی داده‌های حاضر مربوط به دو سال گذشته است.

۲. هنگامی که کاربر به داده‌هایی در دوره‌ای خاص نیاز دارد، ولی داده‌های حاضر مربوط به دوره‌ای دیگر است.

در مدیریت به‌هنگام بودن داده‌ها بر این نکته تأکید می‌شود که باید داده‌ها را زودتر از تاریخ فعلی‌شان منتشر کرد. این امر فقط کمک به کاربران نیست، بلکه سبب ایجاد رویکردی می‌شود که موجب بهبود تصمیم‌ها، حفظ چارچوب کیفیت و افزایش بهره‌وری می‌شود. لازم به ذکر است که با کسب موفقیت در انتشار به‌هنگام داده‌ها، اندازه‌گیری عملکرد مطلوب‌تر خواهد شد. به‌هنگام بودن به حداقل تأخیر برای در

1. timeliness

دسترس قرار دادن داده‌ها می‌پردازد. متغیر به‌هنگام بودن داده‌ها از ارتباط داده‌ها با نیاز سازمان نشأت می‌گیرد. در مدیریت به‌هنگام بودن به این سؤال پاسخ داده می‌شود که داده‌ها بعد از پایان دوره مرجع، با چه فاصله زمانی یا با چه سرعتی برای اطلاع‌رسانی یا پردازش‌های بعدی در دسترس قرار می‌گیرد؟ پاسخ دادن به سؤال بسته به نرخ تغییرات پدیده‌ها و فراوانی آن‌ها دارد (Statistics Canada's Quality Assurance Framework 2002, 15).

۲-۷. تضمین شفافیت و دسترسی^۱

شفافیت به معنای ایجاد شرایط و تمهیداتی است که کاربران آماری بتوانند به‌آسانی انبوهی از داده‌ها و اطلاعات آماری و مفاهیم و تعاریف آن‌ها را درک و فهم کنند (NQAF Glossary 2012, 5). مراکز آماری باید این اطمینان را حاصل کنند که داده‌های مورد نیاز سازمان‌ها را تولید و بدون دشواری در اختیار آن‌ها قرار خواهند داد. این عنصر بر قابلیت دسترسی داده‌های باکیفیت اشاره دارد. همچنین، آن‌ها باید داده‌ها را به‌صورت قابل فهم و ساده و به دور از تکلف منتشر کنند. عدالت انتشار اطلاعات بحث جدی و مهمی است که در این عنصر مطرح می‌شود. مراکز ملی آمار به‌عنوان متولیان انتشار اطلاعات، باید اجازه دسترسی به داده‌های کلان و خرد را برای اهداف تحقیقاتی نیز میسر سازند. در حقیقت این امر باید جزئی از سیاست‌های مراکز باشد (NQAF Glossary 2012, 42).

همچنین، شفافیت و روشنی داده‌های آماری به میزان قابل توجهی بستگی به نحوه انتشار آن‌ها دارد. ارائه خلاصه‌ای از گزارش‌های آماری می‌تواند در کیفیت گزارش‌ها مؤثر باشد. انتشار گزارش‌های آماری در قالب وب موجب سهولت در دسترسی به اطلاعات شده و برای کاربران نسبت به فرمت‌های چاپی آماری ترجیح دارد. در بسیاری از کاربران حرفه‌ای این تمایل وجود دارد که به پایگاه‌های اینترنتی مراجعه کنند؛ چرا که امکان جست‌وجوی داده‌های مورد نیاز آنان وجود دارد که بر اساس آن بتوانند تحلیل خود را ارائه کنند (ESS Standard for Quality Reports 2009, 20).

1. accessibility and clarity

۲-۸. تضمین انسجام^۱

ایجاد شایستگی لازم برای داده‌های آماری تا در مواقع نیاز بتوان آن‌ها را ترکیب و ممزوج نمود و همچنین، کاربران و متقاضیان در زمینه‌های متفاوت بتوانند از آن استفاده کنند (NQAF Glossary 2012, 5). مراکز ملی آمار باید داده‌هایی را منتشر کنند که از قابلیت انسجام درونی برخوردار باشند؛ حتی اگر داده‌های مورد نیاز از چند منبع متفاوت به دست آمده باشند (NQAF Glossary 2012, 45). منظور از انسجام آن است که اگر خروجی داده‌های آماری از انسجام درونی برخوردار باشند، ترکیب آن‌ها از روایی لازم برخوردار خواهد بود (ESS Standard for Quality Reports 2009, 21). در توصیفی دیگر برای تضمین انسجام، هدف آن است که اگر داده‌های مورد نظر از منابع متفاوتی جمع‌آوری شدند، یک پدیده را توصیف کنند (Guidelines for measuring statistical quality 2007, 9).

۲-۹. جامعیت^۲

جامعیت، عنصری تضمین‌کننده گستره و عمق داده‌هاست. جامعیت بیان می‌کند که داده باید جامع و مانع باشند (Batini & Scannapieca 2006, 23). عنصر جامعیت در کیفیت داده به آن معناست که هیچ رکورد و داده‌ای نباید از دست برود. در بسیاری از روش‌های پیمایشی جهت جمع‌آوری داده و نمونه، بسیاری از داده‌ها از بین می‌روند. به‌عنوان مثال، در بسیاری از نمونه‌ها، بعضی از موارد بدون پاسخ می‌مانند و رها می‌شوند. در این موارد می‌توان گفت که داده‌ها دچار فقدان کیفیت هستند. در بسیاری از پایگاه‌های داده همانند پایگاه‌های اطلاعاتی مالی، فقدان و از دست رفتن داده می‌تواند موجب بروز مشکلات و پیامدهای نابهنجاری شود که به آسانی قابل جبران نیست. همچنین، در بسیاری از پایگاه‌های دولتی و حکومتی از دست رفتن داده آثار مخربی را به همراه دارد (Herzog, Scheuren & Winkler 2007, 10).

۲-۱۰. تضمین تفسیر داده‌ها^۳

منظور از تفسیر داده‌ها آن است که داده‌ها باید به گونه‌ای باشند که کاربر بتواند آن‌ها را به آسانی درک کرده و آن‌ها را تبدیل به اطلاعات کند. همچنین، کاربر باید

1. coherence

2. completeness

3. interpretability

بتواند آن‌ها را تجزیه و تحلیل نماید تا پاسخگوی نیازهای سازمان باشد. کسب اطلاعات لازم و فهم اطلاعات آماری در پیگیری اهداف زیر حاصل می‌شود:

۱. طبقه‌بندی و مفهوم‌سازی داده‌های جمع‌آوری شده؛
۲. روش مناسب جهت جمع‌آوری و تکمیل داده‌ها؛
۳. اندازه‌گیری دقیق داده‌ها.

مراکز آماری باید تلاش نمایند که تفسیر داده‌های جمع‌آوری شده برای کاربران قابل فهم بوده و در آن اصلاحات تخصصی و مهم به کار نرفته باشند. لازم به ذکر است که به‌هنگام بودن تفسیر داده‌ها از به‌هنگام بودن داده‌ها جداست. ارزیابی کارآمدی و اثربخشی تفسیر داده‌ها از نظر کاربردی نیز نکته‌ای است که باید به آن توجه شود (Statistics Canada's Quality Assurance Framework 2002, 20–21).

جدول ۱. عناصر بُعد مدیریت محصول آماری در مدیریت کیفیت داده

ردیف عنصر	شرح مفهوم عنصر
۱۱	تضمین مناسبت تضمین مناسبت شامل فرایندی است که در آن مشخص می‌شود سازمان چه داده‌هایی و به کدام نوع آن نیاز دارد.
۲	تضمین درستی منظور از درستی داده‌ها، درجه نزدیکی مقادیر و اندازه‌های محاسبه شده یا برآورد شده از یک سنجه به مقدار واقعی آن سنجه است. این بعد در حقیقت خطاهای نمونه‌گیری شده و نشده را دربرمی‌گیرد.
۳	تضمین به‌هنگامی و به‌موقع بودن اطلاعات حاصل از داده‌ها زمانی قابل ارجاع خواهد بود که دوره داده‌ها به تأخیر نیفتاده باشد و همچنین، تاریخ و زمان داده‌ها مشخص باشد.
۴	تضمین وضوح و دسترسی مراکز آماری باید این اطمینان را حاصل کنند که داده‌های مورد نیاز سازمان‌ها را تولید و بدون دشواری در اختیار آن‌ها قرار خواهند داد.
۵	تضمین انسجام مراکز ملی آمار باید داده‌هایی را منتشر کنند که از قابلیت انسجام درونی برخوردار باشند؛ حتی اگر داده‌های مورد نیاز از چند منبع متفاوت به دست آمده باشند.
۶	جامعیت عنصر جامعیت در کیفیت داده به آن معناست که هیچ رکورد و داده‌ای نباید از دست برود.
۷	تضمین تفسیر داده‌ها اگر داده‌های آماری برای کاربر قابل درک نباشد یا به آسانی قابل درک نباشد، داده‌های آماری کیفیت لازم را نخواهند داشت.

۳. پیشینه تجربی

در پژوهش «آذر، خسروانی و جلالی» بر اساس رویکرد سودا^۱، مؤلفه‌های فرایند کیفیت ارقام آماری از جهت‌ها و دیدگاه‌های مختلف برای ساخت‌دهی به مسئله مورد پایش قرار گرفته که مبنای پایش مطالب علمی مستند و مصاحبه‌های عمیق و نیمه‌ساختاریافته با خبرگان صاحب‌نظر در زمینه کیفیت ارقام آماری «مرکز آمار ایران» است. بعد از مشخص شدن نقشه پایش فرایند کیفیت ارقام آماری، سطح‌بندی کیفیت ارقام، اجزای آن و همچنین، ارتباط بین مؤلفه‌ها و عوامل تأثیرگذار بر آن مشخص می‌شود (۱۳۹۴، ۱).

در پژوهش «حائری مهریزی، یگانگی و حسنی» رهنمودهایی برای طراحی یک سامانه هوشمند به منظور بهبود کیفیت داده در تمامی مراحل تولید داده در یک طرح آمارگیری ارائه شد که با استفاده از این رهنمودها ساختار سامانه‌ای هوشمند برای طرح هزینه و درآمد خانوار، که یکی از طرح‌های مهم «مرکز آمار ایران» است، پیشنهاد شده است (۱۳۸۶، ۱).

در پژوهشی که «یزدان‌طلب» انجام داده، هدف، معرفی ابزارهای فناوری اطلاعات، که در زمینه تولید آمار مورد استفاده قرار می‌گیرند، بوده است. در ادامه، به اثربخشی ابزارها در بهبود کیفیت آمار پرداخته شده است. روش تحقیق این پژوهش پیمایشی و به‌صورت پرسشنامه‌ای بوده است. نتایج حاکی از آن است که با احتمال ۹۵ درصد فناوری اطلاعات در بُعد مناسب، درستی و به‌هنگامی تأثیرگذار است (۱۳۸۸، ۵۰).

در پژوهش «سوکومار، ناتاراجان و فرل» توسعه مدل ارزیابی کیفیت داده بر اساس تجزیه و تحلیل ابر داده‌ها بود. روش این پژوهش با استفاده از مطالعات کتابخانه‌ای و مطالعه موردی بوده است. یافته‌ها حاکی از آن است که ابر داده‌ها در بردارنده اطلاعات مهمی از یک پایگاه داده‌ای هستند و می‌توانند برای ارزیابی کیفیت داده‌ها به تصمیم‌گیرندگان کمک کنند (Sukumar, Natarajan & Ferrell 2015, 621).

در پژوهشی که توسط «مزانزائیکا» و همکاران صورت گرفت، هدف، تحلیل استواری کیفیت داده‌های چندبعدی بود. روش مورد استفاده در این پژوهش بررسی مدل‌های کیفیت داده و در ادامه، به کارگرفتن آن‌ها در یک پایگاه داده با ساختاری ضعیف بوده است. پایگاه‌های داده دربرگیرنده مشاغل کار یک میلیون نفر از مردم بوده است. نتایج،

1. Strategic Option Development and Analysis (SODA)

حکایت از آن داشت که مدل مطلوب در مدیریت کیفیت داده، که دارای دو قسم پایگاه داده‌ها و روش‌های غربالگری است، از اثربخشی بیشتری نزد تصمیم‌گیرندگان برخوردار است (Mezzanzanica et al. 2015, 144).

جدول ۲. پیشینه تجربی پژوهش

نویسندگان	سال	موضوع محوری
آذر، خسروانی، و جلالی	۱۳۹۴	ساخت‌دهی مؤلفه‌های فرایندهای کیفیت اقلام آماری
حائری مهریزی، یگانگی و حسنی	۱۳۸۶	ارائه رهنمودهایی برای طراحی یک سامانه هوشمند به‌منظور بهبود کیفیت داده
یزدان‌طلب	۸۸۱۳	معرفی ابزارهای فناوری اطلاعات در زمینه تولید آمار
Sukumar, Natarajan & Ferrell	۲۰۱۵	توسعه مدل ارزیابی کیفیت داده بر اساس تجزیه و تحلیل ابرداده‌ها
Mezzanzanica et al.	۲۰۱۵	تحلیل استواری کیفیت داده‌های چندبعدی

از بررسی ادبیات حوزه مدیریت کیفیت داده، به حیاتی بودن نقش کیفیت داده‌ها در کلیه عرصه‌های سازمان به‌ویژه در حوزه تصمیم‌گیری مدیران پی برده خواهد شد. پیچیدگی‌های محیطی، سرعت تغییرات، کمبود منابع و سرمایه‌ها و ... همگی از برجسته‌بودن نقش کیفیت داده‌ها و آمارها در عرصه مدیریت کشور خبر می‌دهند. ضمن آن‌که هرچند در پژوهش‌های گذشته به مسئله کیفیت داده پرداخته شده است، اما ابعاد و مؤلفه‌های مدیریت کیفیت داده مورد غفلت قرار گرفته و پژوهشی مدون پیرامون تحلیل عناصر آن انجام نشده است. این امر موجب آن شده است که موضوع کیفیت داده‌ها به مسئله‌ای انتزاعی تبدیل شده که در نتیجه آن مدیران نتوانند به‌طور شایسته به پیاده‌سازی اصول چارچوب ملی تضمین کیفیت داده در سازمان‌های خوب بپردازند. لذا در ادامه، با هدف ارائه راهکارهای کاربردی، به تبیین شدت اثر هر یک از عناصر مدیریت محصول آماری در مدیریت کیفیت داده پرداخته خواهد شد.

۴. روش تحقیق

در پژوهش حاضر، مفاهیم نظری یعنی ابعاد مدیریت کیفیت داده و عناصر بعد مدیریت محصول از طریق مطالعات کتابخانه‌ای جمع‌آوری شد. اصلی‌ترین ابزار گردآوری اطلاعات در پژوهش حاضر، پرسشنامه است. لازم به ذکر است که به دلیل تجربه‌محور

بودن این روش، نیازی به بررسی روایی و پایایی پرسشنامه نبوده است. در مرحله اول با مطالعات کتابخانه‌ای گسترده و استفاده از خبرگان دانشگاهی در گروه کانونی عناصر مدیریت محصول آماری شناسایی شدند. در مرحله دوم، عناصر مدیریت محصول آماری احصاء شده در پرسشنامه‌های اختصاصی به منظور تعیین شدت اثر این عناصر در اختیار ۱۰ نفر از مدیران «مرکز آمار ایران» (مورد مطالعه) قرار گرفت و اطلاعات مورد نیاز از این مدیران احصاء گردید. مدیران مذکور، همگی از متخصصان کیفیت آماری بودند که تعدادی از اساتید دانشگاه‌ها خبرگی آن‌ها را تأیید کرده‌اند.

۴-۱. تکنیک دیمتل

رویکرد «دیمتل»^۱ برای یافتن و حل مشکلات پیچیده استفاده می‌شود و در سال‌های اخیر در زمینه‌های مختلف به کار رفته است. در رویکرد «دیمتل» شش گام اصلی به شرح زیر وجود دارد.

۱. ساختن ماتریس نظرسنجی از پاسخ‌دهندگان: در گام نخست از هر پاسخ‌دهنده خواسته می‌شود اثر مستقیمی که به نظر وی عنصر ابر عنصر دارد را مشخص کند. این اثر می‌تواند با امتیاز P_{ij} مشخص شود.

۲. ساخت ماتریس تصمیم‌گیری اولیه: این ماتریس در حقیقت از میانگین ساده نظرات پاسخ‌دهندگان در مرحله قبل استخراج می‌شود. ماتریس تصمیم‌گیری اولیه را A می‌نامیم و به صورت $A=[a_{ij}]_{n \times n}$ نشان می‌دهیم. لازم به ذکر است که در ماتریس تصمیم‌گیری n تعداد عناصر است و k تعداد خبرگانی که به پرسشنامه مورد نظر پاسخ داده‌اند

$$a_{ij} = \frac{1}{k} \sum_{k=1}^k P_{ij}$$

۳. محاسبه ماتریس اثر اولیه: ماتریس اثر اولیه D ($D=[d_{ij}]_{n \times n}$) از طریق نرمالایز کردن ماتریس تصمیم اولیه A به دست می‌آید. در این ماتریس همه عناصر روی قطر اصلی برابر با صفر است. ماتریس D با استفاده از روابط زیر به دست می‌آید:

$$D=S.A \quad , \quad S > 0$$

یا، و

$$[d_{ij}]_{n \times n} = S[a_{ij}]_{n \times n} \quad , \quad S > 0,$$

$$S = \text{Min} \left[\frac{1}{\sum_{j=1}^n |a_{ij}|}, \frac{1}{\sum_{i=1}^n |a_{ij}|} \right], 1 \leq i \leq n, 1 \leq j \leq n$$

۴. استخراج ماتریس کامل اثر مستقیم و غیرمستقیم: چنانچه توان ماتریس اثر اولیه (D)، افزایش یابد، اثرات مستقیم مسئله کاهش پیدا می‌کند که این امر راه‌حل‌های همگرا به ماتریس معکوس را تضمین می‌کند. ماتریس اثر کل، که T نامیده می‌شود، از رابطه زیر به دست می‌آید:

$$T = D + D^2 + D^3 + \dots + D^m = D(I - D)^{-1}, \quad m \rightarrow \infty$$

اگر مجموع سطرها و ستون‌ها در ماتریس T، به ترتیب، با بردار r و d نمایش داده شوند، خواهیم داشت:

$$T = [t_{ij}]_{n \times n}$$

$$R = [r_i]_{n \times 1} = \left(\sum_{j=1}^n t_{ij} \right)_{n \times 1}$$

$$D = [d_j]_{1 \times n} = \left(\sum_{i=1}^n t_{ij} \right)_{1 \times n}$$

اگر r_i نشان‌دهنده جمع سطری ردیف i ام ماتریس T باشد، پس r_i نشان‌دهنده مجموع اثرات مستقیم و غیرمستقیم عامل i ام بر روی عوامل (معیارها) است. اگر d_j جمع ستونی ستون j ام ماتریس T باشد، پس d_j نشان‌دهنده جمع اثرات مستقیم و غیرمستقیمی است که عامل j ام از دیگر عوامل می‌پذیرد. وقتی که $i=j$ باشد، بدین معناست که جمع سطر و جمع ستون $(r_i + d_j)$ نمایانگر شاخص نشان‌دهنده قدرت اثرپذیری و اثرگذاری است. به بیان دیگر، $r_i + d_j$ نشان می‌دهد که درجه نقش مرکزی که عنصر i در این مسئله بازی می‌کند، چقدر است. بنابراین، اگر $(r_i - d_j)$ مثبت باشد، بیانگر این است که عنصر i بر دیگر عناصر اثر می‌گذارد و اگر منفی باشد، نشان می‌دهد که عنصر i از دیگر عناصر اثر می‌پذیرد.

۵. تعیین ارزش آستانه و به دست آوردن نقشه ارتباط - اثر

با تعیین ارزش آستانه لازم است که ساختار ارتباطی بین عناصر توضیح داده شود. بر اساس ماتریس T هر عنصر t_{ij} از ماتریس T اطلاعاتی در مورد چگونگی تأثیر عنصر i بر عنصر j را توضیح می‌دهد. جهت کاهش پیچیدگی نقشه اثر - ارتباط (IRM)، تصمیم‌گیرندگان باید یک ارزش آستانه برای سطوح اثرگذاری تعیین کنند. بر این اساس، فقط عناصری که سطح تأثیر آن‌ها در ماتریس T از ارزش آستانه بیشتر است، می‌توانند انتخاب شوند و در

IRM نمایش داده شوند. در این ماتریس روی محور افقی (x) ، $r_{i,d}$ و روی محور عمودی (y) ، $r_{i,d}$ قرار می‌گیرند. ارزش‌های $r_{i,d}$ اهمیت هر عامل را نشان می‌دهد و هرچه عاملی مقادیر بالاتری از این ارزش را به خود اختصاص دهد، از اهمیت بالاتری نیز برخوردار است (آذر، خسروانی و جلالی ۱۳۹۴، ۳۲۳-۳۲۷).

۵. تجزیه و تحلیل

در پژوهش حاضر به منظور بررسی شدت اثرگذاری و اثرپذیری عناصر محصول آماری در مدیریت کیفیت داده از تکنیک «دیمتل» استفاده شده است. در مرحله اول، به تهیه پرسشنامه مقایسات زوجی پرداخته شد. این پرسشنامه حاوی ماتریسی است که سطر و ستون آن را عناصر بعد مدیریت محصول آماری (جدول شماره ۱) تشکیل می‌دهند. این جدول در بین ۱۰ نفر از مدیران «مرکز آمار ایران» توزیع شد که حاصل آن در ماتریس تصمیم‌گیری اولیه (A) به شرح زیر است.

جدول ۳. ماتریس تصمیم‌گیری اولیه (A)

عناصر محصول آماری	تضمین تناسب	تضمین درستی	تضمین به هنگامی و به موقع بودن	تضمین وضوح و دسترسی	تضمین انسجام	جامعیت	تضمین تفسیر داده‌ها
تضمین تناسب	۰	۱/۵	۲/۱	۲	۱/۹	۲/۱	۲/۳
تضمین درستی	۱/۳	۰	۲	۲/۸	۳/۳	۳/۲	۲/۸
تضمین به هنگامی و به موقع بودن	۱/۷	۲/۲	۰	۲/۱	۱/۶	۱/۴	۱/۴
تضمین وضوح و دسترسی	۲/۳	۲/۱	۲/۷	۰	۱/۵	۲	۲/۹
تضمین انسجام	۱/۸	۳/۱	۱/۷	۲	۰	۳/۶	۲/۴
جامعیت	۲	۳/۲	۲	۲	۳/۳	۰	۲/۲
تضمین تفسیر داده‌ها	۲/۳	۲/۸	۱/۹	۲/۲	۱/۷	۲	۰

مرحله دوم: در این گام بر اساس فرمولی که در قسمت قبل ارائه شده، به نرمال‌سازی ماتریس تصمیم اولیه (A) و تشکیل ماتریس اثر اولیه (D)، می‌پردازیم که در جدول ۳ این ماتریس نشان داده شده است.

مرحله سوم: در این مرحله با ضرب ماتریس نرمال در معکوس ماتریس یک‌همنه‌ای ماتریس نرمال به ماتریس ارتباط کامل می‌رسیم که در جدول ۴ این ماتریس نشان داده شده است.

جدول ۴. ماتریس اثر اولیه (D)

عناصر محصول آماری	تضمین تناسب	تضمین درستی	تضمین به هنگامی و به موقع بودن	تضمین وضوح و دسترسی	تضمین انسجام	جامعیت	تضمین تفسیر داده‌ها
تضمین تناسب	۰	۰/۰۹۷۴۰۳	۰/۱۳۶۳۶۴	۰/۱۲۹۸۷	۰/۱۲۳۳۷۷	۰/۱۳۶۳۶۴	۰/۱۴۹۳۵۱
تضمین درستی	۰/۰۸۴۴۱۶	۰	۰/۱۲۹۸۷	۰/۱۸۱۸۱۸	۰/۲۱۴۲۸۶	۰/۲۰۷۷۹۲	۰/۱۸۱۸۱۸
تضمین به هنگامی و به موقع بودن	۰/۱۱۰۳۹	۰/۱۴۲۸۵۷	۰	۰/۱۳۶۳۶۴	۰/۱۰۳۸۹۶	۰/۰۹۰۹۰۹	۰/۰۹۰۹۰۹
تضمین وضوح و دسترسی	۰/۱۴۹۳۵۱	۰/۱۳۶۳۶۴	۰/۱۷۵۳۲۵	۰	۰/۰۹۷۴۰۳	۰/۱۲۹۸۷	۰/۱۸۸۳۱۲
تضمین انسجام	۰/۱۱۶۸۸۳	۰/۲۰۱۲۹۹	۰/۱۱۰۳۹	۰/۱۲۹۸۷	۰	۰/۲۳۳۷۶۶	۰/۱۵۵۸۴۴
جامعیت	۰/۱۲۹۸۷	۰/۲۰۷۷۹۲	۰/۱۲۹۸۷	۰/۱۲۹۸۷	۰/۲۱۴۲۸۶	۰	۰/۲۸۵۷۱۴
تضمین تفسیر داده‌ها	۰/۱۴۹۳۵۱	۰/۱۸۱۸۱۸	۰/۱۲۳۳۷۷	۰/۱۴۲۸۵۷	۰/۱۱۰۳۹	۰/۱۲۹۸۷	۰

جدول ۵. ماتریس ارتباط کامل (T)

عناصر محصول آماری	تضمین تناسب	تضمین درستی	تضمین به هنگامی و به موقع بودن	تضمین وضوح و دسترسی	تضمین انسجام	جامعیت	تضمین تفسیر داده‌ها
تضمین تناسب	۰/۶۶۸۴۵۶	۰/۹۴۵۵۸۱	۰/۸۳۹۰۰۸	۰/۸۷۳۹۴۵	۰/۸۸۵۸۴۵	۰/۹۴۲۲۳۲	۰/۹۳۴۶۱۴
تضمین درستی	۰/۹۳۷۳۶۳	۱/۱۰۴۷۱۱	۱/۰۳۷۱۲۵	۱/۱۲۹۸۹۷	۱/۱۸۰۹۱۹	۱/۲۳۶۸۵۱	۱/۱۹۰۴۵۹
تضمین به هنگامی و به موقع بودن	۰/۶۸۵۶۳۲	۰/۸۸۸۳۴	۰/۶۴۴۱۰۱	۰/۸۰۰۶۵۳	۰/۷۹۰۲	۰/۸۲۲۳۲۵	۱/۸۰۶۲۶۸
تضمین وضوح و دسترسی	۰/۸۷۱۲۶۱	۱/۰۶۷۰۹	۰/۹۴۸۲۷۹	۰/۸۴۳۲۰۳	۰/۹۴۹۹۲۵	۱/۰۲۴۷۹۳	۱/۲۰۵۱۸۸
تضمین انسجام	۰/۹۲۵۲۶۳	۱/۲۲۷۲۶۴	۰/۹۸۳۸۴۶	۱/۰۵۱۶۶۳	۰/۹۶۶۴۰۷	۱/۲۱۳۴۵۹	۱/۱۲۸۶۱۲
جامعیت	۰/۹۳۵۱۴۹	۱/۲۳۱۵۶۱	۰/۹۹۹۱۸۷	۱/۰۵۲۸۳۶	۱/۱۴۲۸۹۴	۱/۰۲۳۶۰۴	۱/۱۱۹۸۸
تضمین تفسیر داده‌ها	۰/۸۵۳۴۹۲	۱/۰۸۱۳۵۹	۰/۸۹۰۵۵۷	۰/۹۵۰۸۸	۰/۹۴۵۷۰۸	۱/۰۱۰۲۱۸	۱/۸۷۵۸۰۱

مرحله چهارم: در این مرحله بر اساس میزان تأثیرگذاری و تأثیرپذیری که به ترتیب از جمع سطرها و ستون‌های هر عنصر محاسبه می‌شود، می‌توان به یک نمودار علی و رتبه‌بندی عناصر مدیریت محصول آماری رسید.

جدول ۶. شدت اثر عناصر بر روی یکدیگر

ردیف	عناصر محصول آماری	R	D	D-R	D+R
۱	تضمین مناسبت	۶/۰۸۹۶۸۱	۵/۸۸۶۶۱۷	۰/۲۰۳۰۶۴	۱۱/۹۷۶۳
۲	تضمین درستی	۷/۸۱۷۳۲۵	۷/۵۴۵۹۰۷	۰/۲۷۱۴۱۸	۱۵/۳۶۳۲۳
۳	تضمین به هنگامی و به موقع بودن	۵/۴۴۷۵۱۹	۶/۳۴۲۱۰۳	۰/۸۹۴۵۸-	۱۱/۷۸۹۶۲
۴	تضمین وضوح و دسترسی	۶/۷۵۶۴۳۴	۶/۷۰۳۰۷۷	۰/۰۵۳۳۵۷	۱۳/۴۵۹۵۱
۵	تضمین انسجام	۷/۴۹۶۵۱۳	۶/۸۶۱۸۹۷	۰/۶۳۴۶۱۷	۱۴/۳۵۸۴۱
۶	جامعیت	۷/۵۰۵۱۱۲	۷/۲۷۳۴۸۳	۰/۲۳۱۶۲۸	۱۴/۷۷۸۵۹
۷	تضمین تفسیر داده‌ها	۶/۶۰۸۰۱۴	۷/۱۰۷۵۱۵	۰/۴۹۹۵-	۱۳/۷۱۵۵۳

مرحله پنجم: برای درک بهتر اثر عناصر بر روی یکدیگر نمودار IRM ترسیم می‌شود که در ادامه به آن اشاره می‌شود. به این منظور، مقدار حد آستانه مشخص می‌شود تا روابط با اثرات کم را فیلتر کند. به بیان دیگر، فقط اثراتی نمایش داده می‌شود که مقدار آن‌ها در ماتریس T از حد آستانه بیشتر باشد. در پژوهش حاضر، مقدار حد آستانه عناصر برابر ۰/۹۷ تعیین شده است. در جدول ۷، که همان ماتریس T است، اثرات بیشتر از حد آستانه مشخص شده است.

جدول ۷. تعیین اثرات بیشتر از حد آستانه در ماتریس اثر کل

عناصر محصول آماری	تضمین مناسبت	تضمین درستی	تضمین به هنگامی و به موقع بودن	تضمین وضوح و دسترسی	تضمین انسجام	جامعیت	تضمین تفسیر داده‌ها
تضمین مناسبت	۰.۶۶۸۴۵۶	۰.۹۴۵۵۸۱	۰.۸۳۹۰۰۸	۰.۸۷۳۹۴۵	۰.۸۸۵۸۴۵	۰.۹۴۲۳۳۲	۰.۹۳۴۶۱۴
تضمین درستی	۰.۹۳۷۳۶۳	۱.۱۰۴۷۱۱	۱.۰۳۷۱۲۵	۱.۱۳۹۸۹۷	۱.۱۸۰۹۱۹	۱.۲۳۴۵۶۱	۱.۱۹۰۴۵۹
تضمین به هنگامی و به موقع بودن	۰.۶۹۵۴۳۲	۰.۸۸۸۳۴	۰.۶۴۴۱۰۱	۰.۸۰۰۶۵۳	۰.۷۹۰۲	۰.۸۲۳۳۳۵	۰.۸۰۶۲۶۸
تضمین وضوح و دسترسی	۰.۸۷۱۲۶۱	۱.۰۶۷۰۹	۰.۹۴۸۲۷۹	۰.۸۴۲۲۰۳	۰.۹۴۹۹۲۵	۱.۰۲۴۷۹۳	۱.۰۵۱۸۸۲
تضمین انسجام	۰.۹۲۵۲۶۳	۱.۲۲۷۲۶۴	۰.۹۸۳۸۶۶	۱.۰۵۱۶۶۳	۰.۹۶۶۰۷	۱.۲۱۳۴۵۹	۱.۱۲۸۶۱۲
جامعیت	۰.۹۳۵۱۴۹	۱.۲۳۱۵۶۱	۰.۹۹۹۱۸۷	۱.۰۵۲۸۳۶	۱.۱۴۲۸۹۶	۱.۰۲۳۰۰۴	۱.۱۱۹۸۸
تضمین تفسیر داده‌ها	۰.۸۵۳۴۹۲	۱.۰۸۱۳۵۹	۰.۸۹۰۵۵۷	۰.۹۵۰۸۸	۰.۹۴۵۷۰۸	۱.۰۰۱۰۲۱۸	۰.۸۷۵۸۰۱

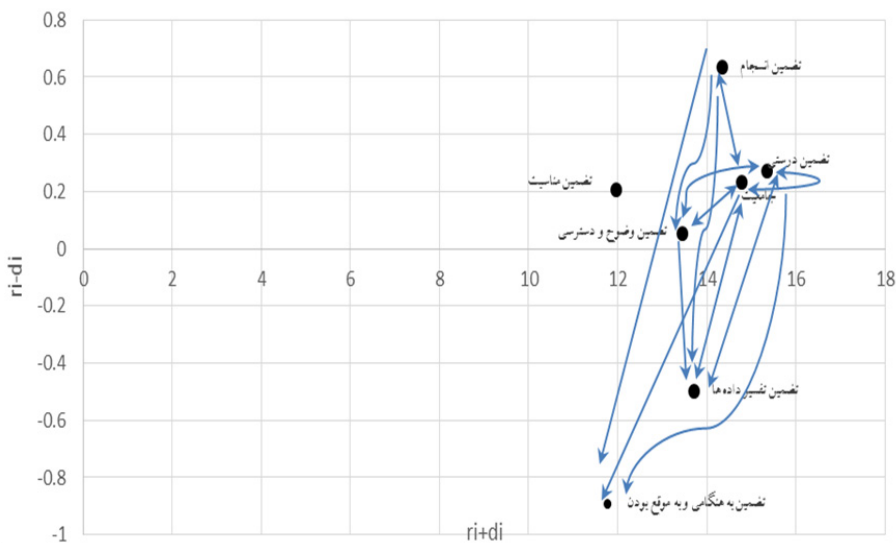
۶. نتایج

همان‌گونه که اشاره شد، امروزه داده‌ها در هر سازمان حکم آب را دارند و اگر از کیفیت لازم برخوردار نباشند بر کلیه تصمیمات سازمان اثر گذاشته و موجب اتلاف فرصت‌ها، هزینه‌های سنگین و در نهایت، شکست سازمان خواهد شد. اما گستردگی و وسعت نظری موضوع مدیریت کیفیت داده و از سوی دیگر، محدودیت‌هایی که مدیران سازمان‌ها با آن مواجه هستند، همانند فشارهای محیطی، محدودیت‌های مالی، محدودیت‌های زمانی، کمبود متخصصان آماری و ... ممکن است نتواند تحقق این امر را ممکن سازد. با وجود این، سازمان‌هایی همچون «مرکز آمار»، «بانک مرکزی»، «سازمان بورس و اوراق بهادار» و ... نیاز مبرمی به پیاده‌سازی عناصر مدیریت کیفیت داده دارند. یکی از عواملی که موجب تسهیل این مهم می‌شود، شناسایی عناصر و تبیین شدت اثر هر یک از آنهاست. بدین صورت، مدیران می‌توانند با سرمایه‌گذاری بر روی عناصر کلیدی و مد نظر قراردادن آنها در برنامه‌ریزی‌های سازمان، موجب ارتقا و تضمین کیفیت داده‌ها و اطلاعات شوند. در پژوهش حاضر با شناسایی عناصر بُعد مدیریت محصول آماری در مدیریت کیفیت داده و اجرای تکنیک «دیمتل» جهت تبیین شدت اثر هر یک از آنها نتایج زیر حاصل شده است:

۱. عناصر جامعیت، تضمین انسجام و تضمین درستی بر کلیه عناصر مدیریت محصول آماری از شدت اثرگذاری بالایی برخوردار بوده و در مراکز آماری باید مبنای برنامه‌ریزی قرار گیرند؛ چرا که منطبق بر فرض «دیمتل» در عناصر مذکور $R+D > M$ و $R-D > 0$ است؛
۲. عناصر تضمین تفسیر داده‌ها از شدت اثرپذیری بالایی برخوردار است. به عبارت دیگر، این عنصر مسئله اصلی بوده و برنامه‌ریزی‌ها باید در جهت حل آن به کار گرفته شوند؛
۳. عنصر تضمین وضوح و دسترسی بر عناصر تضمین درستی، جامعیت و تضمین تفسیر داده‌ها اثر می‌گذارد؛
۴. همچنین، عنصر تضمین تفسیر داده‌ها بر عناصر تضمین درستی و جامعیت اثرگذار است.

لازم به ذکر است که بر اساس نتایج حاصله، عنصر تضمین مناسبت یک عنصر مستقل

است و طبق فرض «دیمتل» که $R-D > 0$ و $R+D = \varepsilon$ است، بر سایر عناصر تأثیر ناچیزی دارد. همچنین، عنصر به‌هنگامی نیز یک عنصر مستقل است که بنابر فرض «دیمتل» که $R-D < 0$ و $R+D = \varepsilon$ است، سایر عناصر بر آن تأثیر می‌گذارند. روابط مذکور در شکل ۲، نشان داده شده است.



شکل ۲. نمودار ارتباط-اثر عناصر مدیریت محصول آماری

بنابراین، عناصر جامعیت، تضمین درستی، و تضمین انسجام از شدت اثر بالایی برخوردار بوده و مدیران مراکز آماری باید برنامه‌های سازمان را در جهت تحقق عناصر مذکور به کار ببندند؛ چرا که تحقق این عناصر علاوه بر افزایش سطح کیفیت داده‌ها موجب اثرگذاری بر سایر عناصر شده و موجب تحقق عناصر دیگر نیز می‌شوند.

۷. محدودیت‌ها و پیشنهادات

پژوهش حاضر با محدودیت‌هایی مواجه بوده است که در ادامه به آن‌ها اشاره خواهد شد؛ ضمن آن که محققان پژوهش حاضر برای سایر پژوهشگرانی که قصد تحقیق در این زمینه را دارند، پیشنهاداتی دارند.

۱. در پژوهش حاضر، به دلیل محدودیت‌های زمانی تنها به ۱۰ نفر از مدیران «مرکز آمار ایران» مراجعه شده است.

۲. پژوهشگران تنها توانستند «مرکز آمار ایران» را مورد بررسی قرار دهند.
۳. به دلیل وسعت حوزه مدیریت کیفیت داده، تنها بُعد مدیریت محصول آماری مورد بررسی قرار گرفت.
۴. شایسته است که پژوهشگران علاقه‌مند به حوزه کیفیت داده، این موضوع را در سایر مراکز همچون «بانک مرکزی»، «بورس کالا» و ... مورد بررسی قرار دهند.
۵. پژوهشگران حوزه مدیریت و آمار می‌توانند به بررسی سایر ابعاد مدیریت کیفیت داده همچون مدیریت نظام آماری، مدیریت فرایند آماری و مدیریت محیط آماری بپردازند.
۶. می‌توان از سایر تکنیک‌های تحقیق در عملیات نرم، همچون تحلیل ساختاری-تفسیری، سودا، تحلیل سلسله‌مراتبی^۳ و ... در پژوهش‌های آتی استفاده نمود.

فهرست منابع

- آذر، عادل، فرزانه خسروانی، و رضا جلالی. ۱۳۹۴. تحقیق در عملیات نرم رویکردهای ساختاردهی مسئله. تهران: انتشارات سامان مدیریت صنعتی.
- آذر، عادل، سعید نجفی توانا، و حسین قربانی. ۱۳۹۴. نگاشت نقشه پایش فرایند کیفیت اقلام آماری مرکز آمار ایران با رویکرد تحلیل و توسعه گزینه‌های استراتژیک (سودا)، نشریه پژوهش‌های مدیریت در ایران (۱۹): ۴-۱۹.
- حائری مهریزی، علی‌اصغر، محمدرضا یگانگی و حسین حسنی. ۱۳۸۶. رهنمودهایی برای بهبود کیفیت داده‌ها در طرح‌های آمارگیری با استفاده از سامانه هوشمند. مجله بررسی‌های آمار رسمی ایران، شماره ۷۰: ۲۰-۱.
- دن میونگ لی، آلی شان. ۱۳۸۳. تجربیات کشور کره در ارزیابی کیفیت آماری. ترجمه محمدرضا اناری. ۱۳۸۳. مجله بررسی‌های آمار رسمی ایران ۶۰: ۶۸-۹۵.
- شباک، اشکان. ۱۳۹۲. نظام مدیریت کیفیت آمارهای رسمی از دیدگاه سازمان‌های مهم بین‌المللی، مبانی و مقایسه. مجله بررسی‌های آمار رسمی ایران (۲۴): ۱: ۵۹-۸۶.
- صالحی، نصرت‌الله. ۱۳۸۲. تحلیلی بر تعامل اجزای نظام برنامه‌ریزی و آماری در یک نظام آماری مطلوب. مجله بررسی‌های آمار رسمی ایران شماره ۵۶: ۲-۱۵.
- یزدان‌طلب، مهناز. ۱۳۸۸. نقش فناوری اطلاعات در بهبود کیفیت داده‌های بازار کار. مجله کار و جامعه ۱۵۵: (۴): ۳۳-۶۹.

- Batini, C., & M. Scannapieca. 2006. *Data quality: Concepts, methodologies and techniques*. Carlo Batini, Monica Scannapieca. Data-centric systems and applications. Berlin, New York: Springer.
- ESS Standard for Quality Reports. 2009. Luxembourg: Office for official publication of the European communities.
- Eye, R. 2014. Information quality maturity: Examining the progress of data quality management and identifying a path forward. *ProQuest Dissertations and Theses* 1-120.
- Geisler, S., C. Quix, S. Weber, & M. Jarke. 2016. Ontology-Based Data Quality Management for Data Streams. *Journal of Data and Information Quality* 7 (4): 1–34. doi:10.1145/2968332
- Guidelines for measuring statistical quality. 2007. London: Office for National Statistics.
- Haug, A, J. S. Arlbjørn, F. Zachariassen, & J. Schlichter. 2013. Master data quality barriers: an empirical investigation. *Industrial Management & Data Systems* 113 (2): 234–249.
- Herzog, T. N., F. Scheuren, & W. E. Winkler. 2007. *Data quality and record linkage techniques*. New York, London: Springer.
- Larburu, N, R. Bults, M. van Sinderen, & H. Hermens. 2015. Quality-of-Data Management for Telemedicine Systems. *Procedia Computer Science* 63: 451–458. <https://doi.org/10.1016/j.procs.2015.08.367> (accessed)
- Leonowich-Graham, P. J. 2003. Data quality framework for small companies. *ProQuest Dissertations and Theses* 1-65.
- Mezzanatica, M, R. Boselli, M. Cesarini, & F. Mercurio. 2015. A model-based evaluation of data quality activities in KDD. *Information Processing & Management* 51 (2): 144–166. <https://doi.org/10.1016/j.ipm.2014.07.007> (accessed)
- Morbey, G. 2013. *Data Quality for Decision Makers: A dialog between a board member and a DQ expert*. (2nd ed. 2013). Wiesbaden: Springer Fachmedien Wiesbaden.
- National Quality Assurance Framework. 2012.
- NQAF Glossary. 2012.
- Pollock, S. E. 2012. Data quality rules in the Analytic Health Repository. The University of Utah, ProQuest Dissertations and Theses: 1- 58.
- Statistics Canada's Quality Assurance Framework. 2002.
- Sukumar, S. R., R. Natarajan, & R. K. Ferrell. 2015. Quality of Big Data in health care. *International Journal of Health Care Quality Assurance* 28 (6): 621–634.
- Tayi, G. K, & D. P. Ballou. 1998. Examining data quality. *Communications of the ACM* 41 (2): 54–57.
- Wang, R. Y, & D. M. Strong. 1996. Beyond Accuracy: What Data Quality Means to Data Consumers. *Journal of Management Information Systems* 12 (4): 5–33.
- Wang, Y. R, M. Ziad, & Y. W. Lee. 2001. *Data quality. Kluwer international series on advances in database systems*. Vol. 23. Boston, Mass, London: Kluwer Academic.

مسعود غفاری

متولد سال ۱۳۷۰، دانشجوی دکتری رشته تحقیق در عملیات از دانشگاه علامه طباطبائی است.
یادگیری ماشینی، تحقیق در عملیات نرم، مدل‌سازی، مدیریت کیفیت و تعالی سازمان از جمله علایق پژوهشی ایشان است.



عادل آذر

متولد سال ۱۳۴۵، دارای مدرک دکتری تحقیق در عملیات از دانشگاه تهران است. ایشان هم‌اکنون استاد تمام دانشگاه تربیت مدرس و رئیس دیوان محاسبات است. تصمیم‌گیری کاربردی، تحلیل آماری، روش تحقیق کمی، علم مدیریت فازی و بودجه ریزی عملیاتی از جمله علایق پژوهشی ایشان است.



اشکان شباک

متولد سال ۱۳۵۶، دارای مدرک دکتری آمار کاربردی است. ایشان استادیار و مدیر گروه پردازش داده‌ها و اطلاع‌رسانی در پژوهشکده آمار است. کیفیت آماری، آمار کاربردی، پردازش داده و کیفیت داده از جمله علایق پژوهشی ایشان است.

