

DB3205

苏州市地方标准

DB3205/T 1164—2024

数字政府 市场监管数据质量监测规范

Digital government—Specification for data quality monitoring of
market supervision

2024-12-25 发布

2024-12-31 实施

苏州市市场监督管理局 发布

目 次

前言	II
1 范围	1
2 规范性引用文件	1
3 术语和定义	1
4 基本原则	2
4.1 目标导向原则	2
4.2 闭环处置原则	2
4.3 持续改进原则	2
5 监测方法及方式	2
5.1 监测方法	2
5.2 监测方式	2
6 监测指标	2
6.1 监测指标框架	2
6.2 数据质量指标	3
6.3 工作质量指标	3
7 监测流程及要求	3
7.1 工作流程	3
7.2 方案规划	4
7.3 监测实施	5
7.4 综合评价	6
7.5 结果应用	6
附录 A（资料性） 数据质量监测指标及权重设置	7
附录 B（资料性） 市场主体注册登记核心数据质量规则示例	8
附录 C（资料性） 数据质量评价报告	10
参考文献	11

前 言

本文件按照GB/T 1.1—2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》的规定起草。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别专利的责任。

本文件由苏州市市场监督管理局提出、归口并组织实施。

本文件起草单位：苏州市市场监督管理信息中心、苏州市质量和标准化院、苏州龙石信息科技有限公司。

本文件主要起草人：陈亮亮、苏循进、顾一清、陆志晗、周文渊、戈泽友、练海荣、孙晓宁、杨毓慧、张未来。

数字政府 市场监管数据质量监测规范

1 范围

本文件规定了市场监管数据质量监测的基本原则、监测方法及方式、监测指标、监测流程及要求等相关内容。

本文件适用于市场监管数据采集、汇聚、应用等环节中的质量监测和管理工作。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

GB/T 36344—2018 信息技术 数据质量评价指标

3 术语和定义

GB/T 35295、GB/T 36073、GB/T 36344—2018 界定的以及下列术语和定义适用于本文件。

3.1

数据质量规则 data quality rule

判断数据是否符合数据质量要求的逻辑约束。

3.2

数据质量监测 data quality monitoring

根据数据质量监测指标，对数据及其工作进行判断和分析，识别数据质量问题并进行全流程管理和评价的过程。

3.3

数据责任机构 data responsible organization

负责数据采集和加工工作的组织。

3.4

数据质量监测机构 data quality monitoring organization

负责数据统筹管理和质量监测工作的组织。

3.5

数据更正 data correcting

对已经采集的错误数据字段状态或属性进行修改或更新操作的行为。

3.6

数据质量治理 data quality governance

按照数据质量监测结果，对数据进行更正并逐步提升数据质量的过程。

4 基本原则

4.1 目标导向原则

数据质量监测应围绕业务需求设定质量目标，围绕质量目标开展工作能够更好地识别和解决数据质量问题。

4.2 闭环处置原则

数据质量监测应强调问题的旁路识别、源头处置以及修复后自动关闭的全流程闭环管理，并分析数据质量问题的根本原因，避免问题数据继续产生。

4.3 持续改进原则

数据质量监测不是一次性工作，应根据数据质量监测结果，不断调整数据质量工作的目标，以适应不断变化的业务需求，推动数据质量的持续提升。

5 监测方法及方式

5.1 监测方法

5.1.1 数据质量监测应采用自动化监测的技术方法，通过软件自动分析和判断结果，并保证监测流程可记录、可追溯。

5.1.2 数据质量监测应采用旁路监测的实施方法，将需要监测的数据同步至监测环境进行独立分析，避免对业务系统产生影响，同时满足数据质量监测过程的性能要求。

5.2 监测方式

数据质量监测应支持全量监测、定量监测和抽样监测三种方式：

- a) 全量监测，对涉及的所有数据进行质量监测；
- b) 定量监测，对特定的范围和时间内的数据进行质量监测；
- c) 抽样监测，制定并执行抽样方案，对抽取的数据进行质量监测。

6 监测指标

6.1 监测指标框架

数据质量监测指标框架包括数据质量指标和工作质量指标两个一级指标及若干二级指标，见图1。在数据质量监测工作中，每个一级指标和二级指标可依据实际情况设置权重。

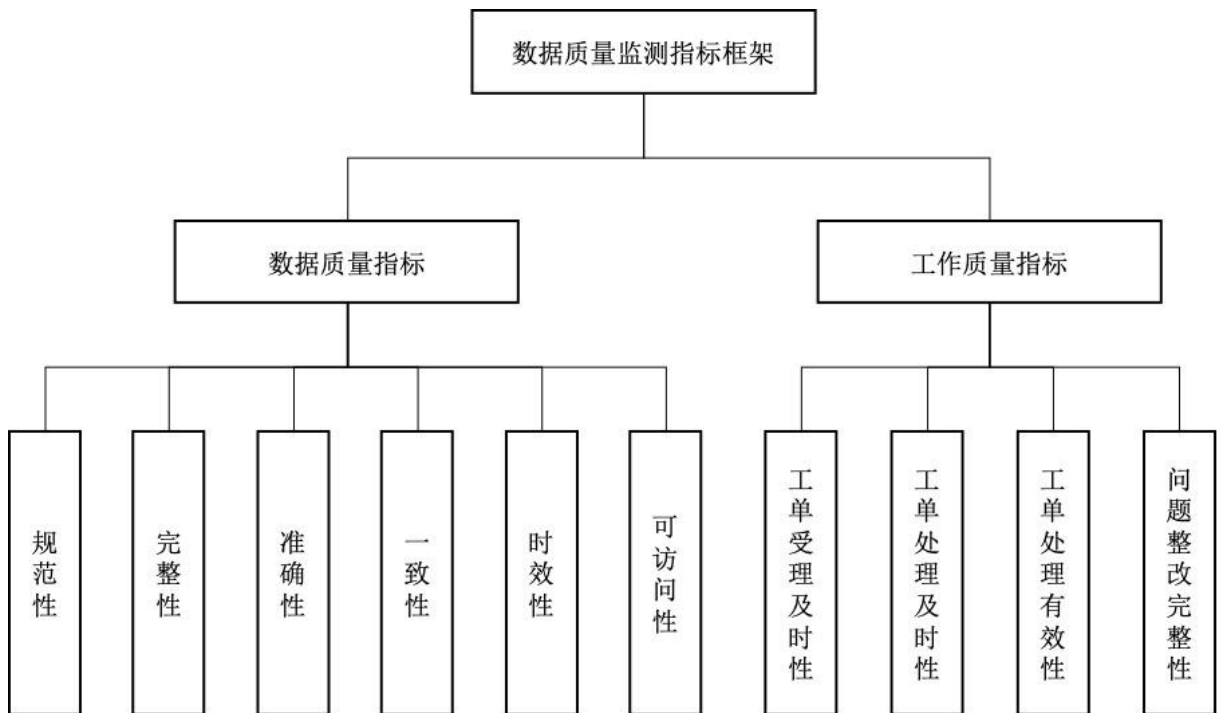


图1 数据质量监测指标框架

6.2 数据质量指标

数据质量指标应符合GB/T 36344—2018中的第4章、第5章的要求，包含规范性、完整性、准确性、一致性、时效性、可访问性六个二级指标：

- a) 规范性：数据符合数据标准、数据模型、业务规则或权威参考数据的程度；
- b) 完整性：按照数据规则要求，记录、属性及字段值等要素的完整程度；
- c) 准确性：表示数据与其所描述实际对象的真实值对比的准确程度；
- d) 一致性：数据与其他特定上下文中使用的数据无矛盾的程度；
- e) 时效性：数据及时有效的程度；
- f) 可访问性：数据能被访问的程度。

注：数据质量指标权重参见附录A。

6.3 工作质量指标

工作质量指标包含工单受理及时性、工单处理及时性、工单处理有效性、问题整改完整性四个二级指标：

- a) 工单受理及时性：在规定时间内，数据责任机构受理工单数与应受理工单数的比率；
- b) 工单处理及时性：在规定时间内，数据责任机构处理工单数与应处理工单数的比率；
- c) 工单处理有效性：数据责任机构处理工单符合整改要求的程度；
- d) 问题整改完整性：在指定时间区间内，数据责任机构问题数据的总体整改率。

注：工作质量指标权重参见附录A。

7 监测流程及要求

7.1 工作流程

数据质量监测工作流程应保证数据质量监测整个过程的有序及有效进行，包括方案制定、数据准备、规则制定、问题发现、问题处置、复测核验、综合评价、结果应用八个环节，见图2。

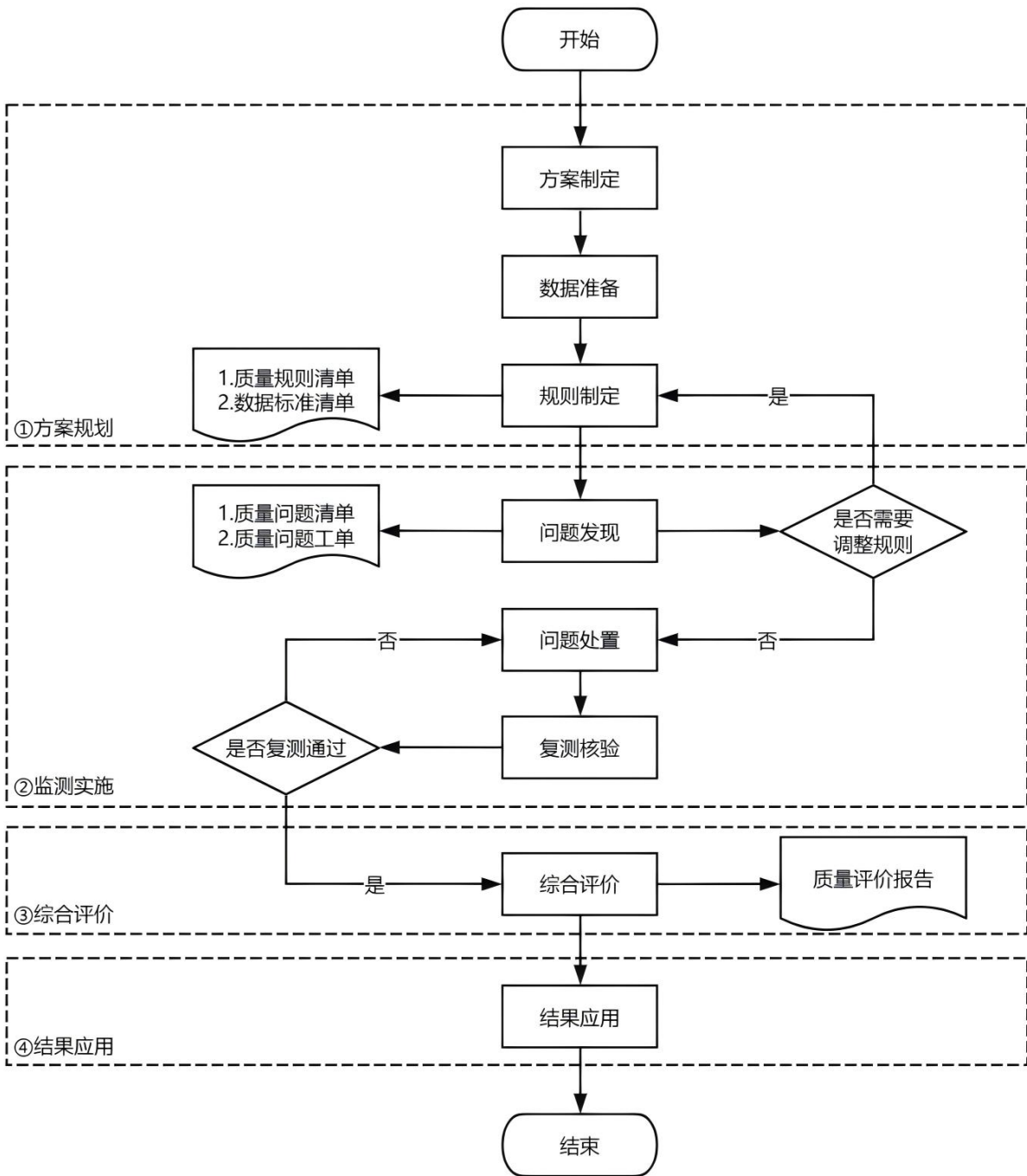


图 2 数据质量监测工作流程

7.2 方案规划

7.2.1 方案制定

7.2.1.1 数据质量监测机构依据实际业务和管理需求，明确数据质量监测的目标。

7.2.1.2 数据质量监测机构依据业务发展的需求以及常见的数据问题的分析，聚焦关键数据和痛点问题，梳理各类数据的优先级，阶段性划定范围。

7.2.1.3 数据质量监测机构制定数据质量监测方案，确认数据范围、监测方式、数据存储方式及位置、质量问题更正要求、实施计划等内容。

7.2.2 数据准备

7.2.2.1 数据质量监测机构与数据责任机构按照监测方案梳理和准备需要监测的数据。

7.2.2.2 数据准备包括一个或多个数据集，以及数据集对应的数据量、元数据描述、更新周期、对应的数据字典信息以及参考数据等内容。

7.2.3 规则制定

7.2.3.1 数据质量监测机构分析数据质量的业务需求和数据要求，识别数据质量特性，收集和确认被监测数据集对应的数据标准及业务规范。

7.2.3.2 数据质量监测机构制定对应的数据质量规则，包括业务规则和技术规则，并确定数据质量监测指标和数据质量规则的权重。

7.2.3.3 数据质量监测机构将数据质量规则进行技术实现，实现数据质量的自动化监测。

注：以市场主体注册登记数据为例，制定的数据质量规则参见附录 B。

7.3 监测实施

7.3.1 问题发现

7.3.1.1 数据质量监测机构根据数据质量监测方案和数据质量规则实施初步监测，发现数据质量问题。

7.3.1.2 数据质量监测机构结合实际业务情况和初步监测结果，验证、优化和调整数据质量规则，直至数据质量监测结果符合预期效果。

7.3.1.3 根据验证的数据质量规则监测数据质量，形成正式的质量问题清单。

7.3.1.4 数据质量监测机构识别数据质量问题的根本原因，包括但不限于信息原因、流程原因、技术原因和人员原因：

- a) 信息原因指数据规范性异常，包括数据标准异常、业务规则异常、变化频率异常等；
- b) 流程原因指信息系统处理流程和人工操作流程设置不当，包括数据的创建、传递、装载、使用、维护等流程；
- c) 技术原因指数据处理各环节的异常，包括业务系统数据校验规则设计不当、数据存储能力有限等，主要涉及数据创建、数据获取、数据传输、数据装载、数据应用、数据维护等环节；
- d) 人员原因指人员及管理机制不足，包括人员培训缺失、责任不明确、激励与反馈机制不足等。

7.3.1.5 数据质量监测机构识别数据质量问题的影响程度，根据数据质量问题的描述和数据应用价值链分析，评估数据质量对业务协同、数据应用等方面的影响，形成数据质量问题影响分析。

7.3.1.6 数据质量监测机构研判的质量问题，研判后进行问题分类和汇总，形成质量问题工单。

7.3.2 问题处置

7.3.2.1 数据质量监测机构将质量问题工单派发至数据责任机构，宜采用线上系统的反馈方式，保障数据安全。

7.3.2.2 数据责任机构在规定时间内接收质量问题工单。

7.3.2.3 数据责任机构在规定时间内处理质量问题，并更新问题工单状态。问题处置状态包括如下几种：

- a) 如确认是数据质量问题，则按照实际情况进行数据更正，问题工单更新为已修复；
- b) 如确认不是数据质量问题，则提供原因说明，问题工单更新为关闭；
- c) 对于短期内无法修复的数据质量问题，则提供原因说明，问题工单更新为暂缓修复。

7.3.2.4 属于数据个案差错的，数据责任机构在源头进行数据更正，并及时向数据质量监测机构提供更正后的正确数据；属于系统性的数据质量问题，应由数据责任机构和数据监测机构共同分析研判，制定并实施更正措施。

7.3.2.5 数据质量监测机构与数据责任机构应保存数据质量问题的更正记录，便于后期追溯和管理。

7.3.3 复测核验

7.3.3.1 数据质量监测机构根据工单的完成情况，对质量问题的整改效果进行核验。

7.3.3.2 针对复测核验不通过的问题工单，由数据质量监测机构发回数据责任机构继续进行数据更正。

7.3.3.3 针对复测核验通过的工单，数据质量监测机构自动关闭问题工单。

7.4 综合评价

7.4.1 数据质量监测机构应定期评价数据质量监测工作的实际成效，发布数据质量评价报告。

7.4.2 数据质量评价报告应遵循科学性、客观性、合理性和全面性。

7.4.3 数据质量评价报告应包括报告范围、评价结果、数据质量规则清单、问题案例清单、监测依据、问题分类等内容。数据质量评价报告要素及说明参见附录 C。

7.4.4 数据质量评价报告中的评价结果应采用可量化的方法客观反映数据质量监测工作的实际情况，具体评价方法参见附录 C。

7.5 结果应用

7.5.1 数据质量监测机构应根据数据质量监测实际结果，制定下一阶段数据质量监测工作的目标。

7.5.2 数据质量监测机构宜收集整理各类数据质量知识、经验和案例，形成数据质量知识库，并通过相关培训、宣讲等活动宣传数据质量监测经验做法，持续提升市场监管数据质量治理能力。

7.5.3 数据责任机构宜根据数据质量评价报告，分析数据质量问题的根本原因，持续提升源头数据质量，进而提高工作与决策的科学性、准确性，为数字政府建设提供支撑。

附 录 A
(资料性)
数据质量监测指标及权重设置

数据质量监测指标及权重设置见表A.1，可根据实际情况调整。

表 A.1 数据质量监测指标及权重设置

序号	一级指标	指标权重	二级指标	指标权重	计算方法
1	数据质量 指标	70	规范性	20	符合规范性规则的数据量/监测数据集的全部数据量
2			完整性	20	符合完整性规则的数据量/监测数据集的全部数据量
3			准确性	30	符合准确性规则的数据量/监测数据集的全部数据量
4			一致性	15	符合一致性规则的数据量/监测数据集的全部数据量
5			时效性	10	符合时效性规则的数据量/监测数据集的全部数据量
6			可访问性	5	符合可访问规则的数据量/监测数据集的全部数据量
7	工作质量 指标	30	工单受理及时性	10	在规定时间内完成受理的工单数量/工单总量
8			工单处理及时性	20	在规定时间内完成处理的工单数据量/工单总量
9			工单处理有效性	40	在规定时间内完成处理的且复测核验通过的工单数量/在规定时间内完成处理的工单总量
10			问题整改完整性	30	在指定时间区间内,完成更正的问题数量/问题总量

附 录 B
(资料性)

市场主体注册登记核心数据质量规则示例

表B.1给出了针对市场监管系统市场主体注册登记核心数据质量规则示例。

表 B.1 市场主体注册登记核心数据质量规则示例

序号	一级指标	二级指标	数据质量规则	评测字段	数据质量规则描述
1	数据质量指标	规范性	统一社会信用代码规范性检查	统一社会信用代码	统一社会信用代码编码规范应该符合行业标准《GB 32100-2015 法人和其他组织统一社会信用代码编码规则》及《GB 32100-2015XG1-2016 法人和其他组织统一社会信用代码编码规则国家标准第1号修改单》。
2			组织机构代码规范性检查	组织机构代码	组织机构代码编码规范应该符合行业标准《GB/T 11714-1997 全国组织机构代码编制规则》。
3			行业分类规范性检查	行业分类	行业分类应符合国家标准《GB/T 4754-2017 国民经济行业分类》。
4			法定代表人或负责人证件号码规范性检查	法定代表人或负责人证件号码	根据证件类型，检测证件号码。如果证件类型为中华人民共和国居民身份证，则证件号码应符合《GB 11643-1999 公民身份号码》。
5			企业状态是否符合既定规范	企业状态	企业状态应符合行业标准《SG 3-2021 市场监管基础代码集》。
6			所属行政区划是否符合相应的标准	所属行政区划	行政区划应符合《GB/T 2260-2007 中华人民共和国行政区划代码》的编码规则，并依据最新区划进行扩展。
7			登记机关规范性检查	登记机关	登记机关应符合行业标准《SG 3-2021 市场监管基础代码集》。
8		完整性	法人基本信息非空检查	机构名称、机构类型、注册地址、经营范围、法定代表人或负责人姓名、法定代表人或负责人证件类型、法定代表人或负责人证件号码不可为空	机构名称、机构类型、注册地址、经营范围、法定代表人或负责人姓名、法定代表人或负责人证件类型、法定代表人或负责人证件号码不可为空。
9			注册资本和注册资本币种相互对应检查	注册资本、注册资本币种	注册资本和注册资本币种需要同时存在。当注册资本有记录时，注册资本币种应当有相应的类型，反之亦然。

表 B.1 市场主体注册登记核心数据质量规则示例（续）

序号	一级指标	二级指标	数据质量规则	评测字段	数据质量规则描述
10	数据质量指标	准确性	经营期限自和经营期限至的逻辑检查	经营期限自、经营期限至	经营（驻在）期限至不得早于经营（驻在）期限自。
11			统一社会信用代码唯一性检查	统一社会信用代码在检查表中不可重复	统一社会信用代码为市场主体唯一标识，不允许重复。
12			成立日期和核准日期逻辑检查	核准日期、成立日期	核准日期不得早于成立日期。
13		一致性	法定代表人或负责人信息与人口基础数据是否一致	法定代表人或负责人姓名、证件号码	法定代表人基本信息与人口基础数据交叉比对，结果是否一致。
14			企业信息的最新数据检查	企业名称、经营范围、住所、法定代表人、市场主体类型、注册资本	企业名称、经营范围、住所、法定代表人、市场主体类型、注册资本与变更备案信息表中最新的变更一致。
15		时效性	数据更新时间是否满足应用需求	数据更新时间	检测核准时间与数据入库时间的差是否在指定区间内。
16	可访问性	所有数据均可正常获取		在指定周期内的市场主体数据均可正常获取。	

附 录 C
(资料性)
数据质量评价报告

C.1 数据质量评价报告的主要内容

表C.1给出了数据质量评价报告的主要内容。

表 C.1 数据质量评价报告的主要内容

序号	章节	是否必选	描述
1	报告范围	必选	报告所监测数据集的范围说明，包括数据集名称、数据量、元数据等内容。
2	评价结果	必选	数据质量监测结果的数字化、量化结果描述，包括总评分、二级指标评分、问题数据在各个指标的分布情况、问题数据在各个数据责任机构的分布情况等内容。
3	数据质量规则清单	必选	数据质量监测所使用的数据质量规则清单。
4	问题案例清单	必选	数据质量监测的问题清单，包括数据标识、质量规则、问题描述、错误原因、严重程度、修复建议等内容。
5	监测依据	必选	数据质量监测的依据，包括国家标准、行业标准、地方标准、团体标准、业务规范等内容。
6	问题分类	可选	根据原因分析，针对所产生的问题从信息、流程、技术、人员等维度的分类。
7	其他	可选	可根据实际情况补充其他章节的内容。
8	数据质量评价方法	可选	数据质量评价方法见附录 C.2。

C.2 数据质量监测量化评价方法

数据质量监测评价方法采用加权平均法计算最终结果。

- a) 根据二级指标下各质量规则的影响程度权重、评测数据量和问题数，计算各项二级指标得分；
- b) 根据各项二级指标得分和权重，计算一级指标得分；
- c) 根据各项一级指标得分和权重，计算数据质量监测最终得分。

参 考 文 献

- [1] GB/T 35295 信息技术 大数据 术语
 - [2] GB/T 36073 数据管理能力成熟度评估模型
 - [3] 《江苏省公共数据管理办法》（省政府令第148号）
 - [4] 《市场监管信息化标准体系》（国市监办发〔2021〕3号）
 - [5] 《关于开展市场主体数据质量提升行动的通知》（市监注发〔2021〕19号）
 - [6] 《关于开展企业信用监管数据质量全面提升行动的通知》（国市监信发〔2023〕25号）
-