

# 团 体 标 准

T/CAFFCI XXXX—202X

## 中国化妆品产业数据统计规范

Group standard of China Cosmetics Industry Data Statistics Code

(征求意见稿)

在提交反馈意见时，请将您知道的相关专利连同支持性文件一并附上。

202X-XX-XX 发布

202X-XX-XX 实施

中国香料香精化妆品工业协会 发布

# 目 次

前 言.....	错误!未定义书签。
引 言.....	错误!未定义书签。
1 范围 .....	错误!未定义书签。
2 规范性引用文件.....	错误!未定义书签。
3 术语和定义.....	错误!未定义书签。
3.1 化妆品 cosmetic.....	错误!未定义书签。
3.2 化妆品产业数据 Cosmetics Industry Data.....	4
3.3 销售渠道 Sales Channel.....	错误!未定义书签。
3.4 商品交易总额 Gross Merchandise Value.....	5
3.5 消费者声量 Consumer Voice Volume.....	5
3.6 净推荐值 Net Promoter Score.....	5
4 计算公式及应用场景.....	5
4.1 商品交易总额计算公式及应用场景 .....	5
4.2 消费者声量计算公式及应用场景.....	6
4.3 净推荐值计算公式及应用场景.....	7
5 数据统计口径.....	8
5.1 统计边界.....	9
5.2 核心统计指标及分类.....	8
5.3 统计原则 .....	9
6 数据的采集、清洗、建仓、呈现和管理.....	7
6.1 多源数据的整合采集.....	9
6.2 数据的清洗.....	错误!未定义书签。
6.3 数据仓库的构建.....	错误!未定义书签。
6.4 数据的可视化呈现.....	错误!未定义书签。
6.5 数据的管理.....	错误!未定义书签。
7 数据分析方法.....	11
7.1 基础分析.....	11
7.2 高级分析.....	11
8 数据发布与共享.....	11
8.1 发布机制.....	11
8.2 共享规范.....	12

# 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》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。

本文件由中国香料香精化妆品工业协会提出并归口。

本文件起草单位：中国香料香精化妆品工业协会、未来派（武汉）信息科技有限公司、青眼网络科技（武汉）有限公司、美妆头条传媒（广州）有限公司、艾媒咨询（广州）有限公司、广州新数智能科技有限公司、上海语析信息咨询有限公司。

本文件主要起草人：姚永斌、刘李军、黄志东。

# 引 言

中国化妆品产业规模已突破万亿元，但当前行业数据的统计存在口径差异、碎片化、标准缺失等问题，导致决策依据不足、国际竞争力受限，制约了行业向高质量发展转型。

本文件将通过三个方面破局：一是统一数据标准，明确化妆品产业生产、流通、研发等环节的关键指标的定义与计算方法；二是融合国际经验与国内实践，借鉴欧美成熟的统计框架，适配直播电商、跨境供应链等中国特色场景；三是构建全链条治理体系，覆盖数据采集、清洗、存储、分析、使用、发布各环节，强化数据真实性与共享安全性。

本文件旨在以标准化驱动行业协同，为中国从“化妆品消费大国”迈向“产业创新强国”提供坚实的数据基石。

# 中国化妆品产业数据统计规范

## 1 范围

本文件规定了化妆品产业在研发、生产、流通、消费等环节进行数据统计时，所涉及的关键指标的定义、计算公式以及数据的采集方法、分析方法、质量控制与分享规范。

本文件适用于化妆品产业数据的采集、处理、分析和分享。

## 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 10092 数据的统计处理和解释 测试结果的多重比较

ISO/IEC 20547-3 信息技术 大数据参考架构 第3部分：参考框架

## 3 术语和定义

下列术语和定义适用于本文件。

### 3.1

#### **化妆品** cosmetic

是指以涂擦、喷洒或者其他类似方法，施用于皮肤、毛发、指甲、口唇等人体表面，以清洁、保护、美化、修饰为目的的日用化学工业产品。为统一口径，方便推动化妆品产业数据的共建与分享，根据现行法律法规，对化妆品的分类作如下说明：

[来源：化妆品监督管理条例 第三条]

### 3.2

#### **化妆品产业数据** cosmetics industry data

指与化妆品研发、生产、流通、消费等相关联的所有结构化与非结构化数据，包括但不限于市场交易数据、产品属性数据、消费者行为数据、供应链数据等。

### 3.3

#### **销售渠道** sales channel

是指商品或服务从生产者向最终消费者转移的路径和网络。

### 3.4

#### 商品交易总额 gross merchandise value

是电商平台或者企业通过电商平台在特定时间段内所有成交订单的总金额。

注 1:商品交易总额的统计起点是订单生成,包含已支付订单、未支付订单(如拍下未付款)、取消订单、拒收订单及退货订单的金额。通常情况下,消费者浏览、收藏或加入购物车等行为所对应的商品或服务金额并不包含在内。

注 2:商品交易总额不扣除退款、折扣、运费等成本,不等于实际收入。是一个衡量市场交易规模和消费者购买意向的非财务指标。

### 3.5

#### 消费者声量 consumer voice volume

是指消费者在社交媒体平台上主动发布的与品牌、商品或服务相关的帖子、评论、转发等内容的总量,反映商品或服务在消费者群体中的存在感、讨论热度和自发传播能力。

注:消费者声量核心在于区分消费者生成内容(UGC)与专业生成内容(PGC),强调消费者自主发声的规模与影响力。

### 3.6

#### 净推荐值 net promoter score

是指衡量消费者忠诚度与推荐意愿的核心指标。

注 1:净推荐值不适用于决策链条长且复杂的 B2B 生意,多用于面向消费者的行为研究。

注 2:其核心提问方式为:“您有多大可能将我们的商品或服务推荐给朋友或同事?”,消费者在 0-10 分的范围内评分。并根据得分分为推荐者(9-10 分,表示高度满意并愿意主动推广商品或服务)、被动者(7-8 分,态度中立,但易受竞争对手影响)、贬损者(0-6 分,对商品或服务不满,可能传播负面口碑)。

## 4 计算公式及应用场景

### 4.1 商品交易总额计算公式及应用场景

#### 4.1.1 商品交易总额计算公式

$$GMV_T = \sum_{i=1}^n (Q_i \times P_i)$$

式中：

$GMV_T$ —时间段 T 内的商品交易总额 (Gross Merchandise Value)，单位为货币；

$Q_i$ —第 i 种商品在时间段T内的销售数量 (Quantity Sold)，单位为件或个；

$P_i$ —第 i 种商品的交易单价 (Price per Unit)，单位为货币/件；

$n$ —时间段 T 内独立商品的总类别数；

$T$ —统计时间范围，如日、月、年等。

#### 4.1.2 商品交易总额应用场景

a) 评估市场地位：商品交易总额大表明消费者基数大、商品种类丰富，有较高的市场地位和增长潜力。

b) 制定营销策略：通过从品类/地区/人群等指标对商品交易总额进行分析，制定和优化营销策略。

c) 判断市场规模：通过横向比较不同电商平台/品牌/企业商品交易总额的大小及增长率，判断市场竞争格局和市场规模。

### 4.2 消费者声量计算公式及应用场景

#### 4.2.1 消费者声量计算公式

消费者声量是数字化时代品牌资产的核心指标之一，其计算需兼顾基础数据与多维加权模型，应用场景覆盖从营销到产品优化的全生命周期。需要结合互动行为（如转发、点赞）对声量的放大效应，以及情感倾向和媒体影响力差异，进行计算。

$$CVV = \sum_{i=1}^n \left( (P_i + C_i) \times (1 + \alpha R_i + \beta L_i) \times S_{e,i} \times W_{c,i} \right)$$

式中：

$CVV$ —某品牌或者商品在一定时间内的消费者声量总额；

$P_i$ —第i个媒体的有效帖子数（去重后）；

$C_i$ —第i个媒体的有效评论数；

$R_i$ —平均每条内容的转发次数；

$L_i$ —平均每条内容的点赞次数；

$\alpha, \beta$ —转发和点赞的权重系数（需通过回归分析或行业基准确定，推荐基准值  $\alpha = 1.5, \beta = 1.0$ ；可根据不同的媒体属性进行调整）；

$S_{e,i}$ —第i个媒体的情感得分（推荐正面=1.2，中性=1，负面=0.8）；

$W_{c,i}$ —第*i*个媒体的传播权重（推荐权重高=1.5~1.0，一般=1.0~0.7，较低=0.7~0.5，低=0.5~0.1）；

*n*—数据采集的媒体总数。

#### 4.2.2 消费者声量应用场景

a) 营销效果评估：通过监测消费者声量，从传播量、品牌美誉度提升等多维度，评估营销效果；并根据不同媒体的传播效果，优化传播策略。

b) 竞品对标与市场定位：通过分析竞品的消费者声量，寻找差距，改进策略。

c) 产品研发与消费者体验优化：通过监测消费者对产品评价的关键词，研发新产品或者改进产品体验。

#### 4.3 净推荐值计算公式及应用场景

##### 4.3.1 净推荐值计算公式

$$NPS = \left( \frac{P - D}{N} \right) \times 100$$

式中：

*NPS*—一定时间内，被调查品牌的净推荐值；

*P*—推荐者（Promoters）数量（评分 9-10 分的消费者）；

*D*—贬损者（Detractors）数量（评分 0-6 分的消费者）；

*N*—总有效样本数（ $N = P + Q + D$ ，*Q*为被动者数量，评分 7-8 分，不计入计算但影响分母）。

结果范围： $-100 \leq NPS \leq +100$ ，数值越高代表消费者忠诚度越高。

##### 4.3.2 净推荐值应用场景

a) 业务预测与增长驱动：测试验证产品口碑后再大规模推广，降低市场风险；通过持续监测净推荐值，实时调整和改进，驱动可持续增长。

b) 产品与服务优化：结合净推荐值的定量得分与定性反馈，企业可定位消费者在购买和使用商品或服务过程中的痛点，从而识别出关键改进点。

### 5 数据统计口径

#### 5.1 统计边界

依据《化妆品监督管理条例》，排除非化妆品类目，建立严谨的细分类目目录。

为统一口径，方便推动化妆品产业数据的共建与分享，根据现行法律法规，化妆品可进

一步细分为如下细分类目：

1) 皮肤护理产品：指作用于皮肤表面，起清洁和护理作用的相关产品。主要包括面部护理与清洁、身体护理与清洁、手/足/胸/臀护理与清洁，防晒、卸妆；

2) 头发护理产品：指作用于头发和头皮，起清洁和护理作用的相关产品。主要包括洗发、护发、染发、造型、烫发相关产品；

3) 彩妆：指作用于面部或者身体表面，起颜色修饰作用的相关产品。主要包括面部彩妆、眼部彩妆、唇部彩妆、底妆、美甲等产品；

4) 香水/香氛：以涂抹、喷洒或其他类似方法，施于人体表面，达到美化、修饰人体气味，通过不同的香气增加人体魅力或吸引力目的的化妆品。

## 5.2 核心统计指标及分类

为统一口径，方便推动化妆品产业数据的共建与分享，化妆品产业数据可进行如下分类：

1) 按数据性质：财务数据（主要包括：销售额、销售量、利润、研发费用、研发费用资本化等）、非财务数据（主要包括：商品交易总额、消费者声量、净推荐值、消费者满意度等）。

2) 按统计层级：宏观数据（全国市场规模、进出口总额等）、中观数据（区域市场表现、品类结构）、微观数据（企业销售数据、品牌和单品销售数据等）。

3) 按更新频率：实时数据、月度/季度数据、年度数据。

## 5.3 统计原则

严格区分不同指标的内涵和外延，选取对应的统计对象，使用科学的统计方法进行统计。

1) 财务数据的统计（销售额、销售量、利润、研发费用、研发费用资本化等）需要符合财政部下发的《企业会计准则》或国际财务报告准则（IFRS）的要求。

2) 非财务数据的统计（消费者声量、商品交易总额等），需要遵循过程严谨科学、数据颗粒度可追溯到最小统计单位、应统计尽统计原则，确保结论的科学性。

3) 涉及全国性的宏观数据统计，需要整合线上线下各销售渠道数据进行全渠道、全口径数据统计。以样本数据为基础，通过数据模型推理全口径总量的必须过拟合并进行交叉验证。

## 6 数据的采集、清洗、建仓、呈现和管理

## 6.1 多源数据的整合采集

进行数据采集之前需要首先明确数据采集标准（如《统计用产品分类目录》），制定采集规范。采集过程中需要通过电子表单校验（如必填项、格式限制）减少录入错误并尽可能采用自动化采集，避免人工填报导致的误差。

1) 官方渠道：政府机关和行业协会等官方机构发布的行业报告和公开数据。

2) 零售商数据：通过 API 接口或数据合作获取电商平台（如天猫、京东等）、线下零售商销售数据。

3) 市场调研数据：问卷调查（净推荐值等）、焦点小组访谈（产品试用反馈等）。抽样调查需满足最低样本要求，定量分析有效样本不低于 1000 个；定性分析有效社会性走访量不低于 20 个；全国性普查样本数不低于 3000。

4) 企业数据：通过征询、申报、走访等方式，从相关企业获取的数据。

5) 第三方数据：采用自动化工具（网络爬虫抓取电商评论等，IoT 设备获取门店客流，POS 机、SASS 收银系统获取销售数据等）获取的数据。第三方市场调研和数据监测公司提供的的数据。

## 6.2 数据的清洗

采用自动化清洗和人工复核双重机制，对数据缺省值、异常值等进行处理。使用专业工具（如 DataStage）对数据进行清洗，主要包括去除噪声数据、填充缺省值、统一数据格式、统一行业术语等，以确保基础数据的质量。经过清洗后的数据需同时满足如下要求：

1) 完整性：覆盖范围全面，数据字段和记录完整；

2) 准确性：多源数据可交叉验证（如平台数据与第三方数据之间进行逻辑校验）；

3) 可靠性：数据来源可靠，采集方法科学，清洗过程操作严谨；

4) 一致性：同一数据在不同系统、时间或表达形式中保持一致。历史数据与当前数据逻辑连贯一致；

5) 时效性：数据能及时反映市场变化，回应市场需求。

## 6.3 数据仓库的构建

依据 ISO/IEC 20547-3 大数据参考架构标准，从用户视角与功能视角双维度出发，构建标准化数据仓库结构、通用功能组件及接口。

## 6.4 数据的可视化呈现

运用专业商业智能及数据分析软件，实现数据的自动化智能分析及可视化呈现，支持多种图表类型和布局方式，确保数据展示的直观性和美观性。

## 6.5 数据的管理

**任务调度：**采用专业工具实现数据仓库任务的自动化调度和监控，确保数据处理流程的高效执行。

**监控与预警：**建立数据仓库的监控体系，实时监控数据处理任务的运行状态，及时发现并解决潜在问题。

**数据质量管理：**制定数据质量评估指标，定期对数据仓库中的数据进行质量检查和优化，确保数据的准确性和完整性。

**文档管理：**规范数据仓库的文档管理，包括数据字典、接口文档、流程图等，确保文档的完整性和可追溯性。

## 7 数据分析方法

### 7.1 基础分析

1) 描述性统计：总额、均值、趋势图、同比、环比、复合增长率等。

2) 关联分析：如通过算法进行消费行为分析，挖掘产品组合规律，分析某一产品的连带率等。

### 7.2 高级分析

1) 聚类分析：将统计对象划分为若干组，使得同一组内的对象具有高度相似性，而不同组间的对象差异显著，进行分析。如按照中国品牌 and 外国品牌进行分类，对比分析各自优劣势等。

2) 推断性分析：如采用时间序列分析 (ARIMA) 预测市场增长率，再结合 GB/T10092-2009 《数据的统计处理和解释——测试结果的多重比较》进行验证。

3) 机器学习：引入 AI 人工智能，通过构建专业知识库进行人工智能训练，然后进行智能化产业数据分析。

## 8 数据发布与共享

### 8.1 发布机制

1) 频率：月度核心指标 (零售总额、同比增长率等)、年度关键数据 (市场趋势分析)。

2) 渠道：行业协会官网及融媒体、国家数据平台，行业媒体、财经媒体、行业会议等。

## 8.2 共享规范

1) 分级共享：脱敏后的宏观数据（如市场规模增长率等）可进行公开传播；涉及国家机密和企业商业机密的数据，以及非公开获取的数据，需要签订保密协议后在允许的范围内传播。

2) 技术保障：差分隐私保护，如涉及保密数据的共享需要进行加密传输、访问日志留存 $\geq 6$ 个月。

3) 规范引用：引用数据需规范注明出处。

---