1. **会员信息基础表hy\_danweiinfo**

|  |  |  |  |
| --- | --- | --- | --- |
| **清洗项** | **清洗规则** | **目的** | **统计结果的影响** |
| 单位名称 | 必填 | 去除无效数据。 | 对大数据统计无影响，可能对业务有影响。 |
| 注册资本 | 大于等于0 | 去除无效数据 | 影响注册资本分布的统计 |
| 统一社会信用代码 | 必填 | 去除无效数据 | 对大数据统计无影响，可能对业务有影响。 |
| 审核时间 | 投标人必填 | 去除无效数据。利于增量抽取。 | 影响今日新增数，影响大数据增量抽取 |
| 小于当前时间，大于某个固定值（默认为1990年） | 去除无效数据。 |

1. **从业人员信息基础表hy\_personinfo**

|  |  |  |  |
| --- | --- | --- | --- |
| **清洗项** | **清洗规则** | **目的** | **统计结果的影响** |
| 人员名称 | 必填 | 去除无效数据，利于人员数的统计。 | 对大数据统计无影响，可能对业务有影响。 |
| 执业资格 | 必填 | 去除无效数据 | 对大数据统计无影响，可能对业务有影响。 |
| 符合“交易平台评委职称类型”代码项 | 去除非法数据 | 对大数据统计无影响，可能对业务有影响。 |
| 审核时间 | 必填 | 去除无效数据。利于增量抽取。 | 影响今日新增数，影响大数据增量抽取 |
| 小于当前时间，大于某个固定值（默认为1990年） | 去除无效数据。 |

1. **企业业绩表hy\_qiyeyejiinfo**

|  |  |  |  |
| --- | --- | --- | --- |
| **清洗项** | **清洗规则** | **目的** | **统计结果的影响** |
| 单位标识 | 必填 | 去除无效数据 | 对大数据统计无影响，可能对业务有影响。也可能在统计业绩时，关联不到相关业绩。 |
| 标段标识 | 必填 | 去除无效数据 | 对大数据统计无影响，可能对业务有影响。也可能在统计业绩时，关联不到相关业绩。 |
| 中标金额 | 大于等于0 | 去除非法数据，利于中标金额总量的统计，排行统计等。 | 对大数据统计暂无影响。 |
| 审核时间 | 必填 | 去除无效数据，利于增量抽取。 | 影响大数据增量抽取。影响大数据增量抽取 |
| 小于当前时间，大于某个固定值（默认为1990年） | 去除无效数据。 |

1. **企业资质详细信息表hy\_qyzizhiinfo**

|  |  |  |  |
| --- | --- | --- | --- |
| **清洗项** | **清洗规则** | **目的** | **统计结果的影响** |
| 单位标识 | 必填 | 去除无效数据 | 影响不同企业资质中标标段数量分布统计 |
| 资质名称 | 必填 | 去除无效数据 |
| 资质代码 | 必填 | 去除无效数据 |
| 审核时间 | 必填 | 去除无效数据，利于增量抽取。 |
| 小于当前时间，大于某个固定值 | 去除无效数据，利于根据时间段的统计。 |

1. **项目信息基础表yw\_projectinfo**

|  |  |  |  |
| --- | --- | --- | --- |
| **清洗项** | **清洗规则** | **目的** | **统计结果的影响** |
| 项目标识 | 必填 | 去除无效数据 | 对大数据统计无影响，可能对业务有影响。也可能在统计业绩时，无法做关联查询。 |
| 项目名称 | 必填 | 去除无效数据 | 对大数据统计无影响，可能对业务有影响。 |
| 项目编号 | 必填 | 去除无效数据 | 对大数据统计无影响，可能对业务有影响。 |
| 项目金额 | 大于等于0 | 去除非法数据，利于项目金额的总数统计，排行等。 | 对大数据统计暂无影响，可能对业务有影响。 |
| 审核时间 | 必填 | 去除无效数据，利于增量抽取。 | 影响本月、本季度、本年等已受理项目数的统计 |
| 小于当前时间，大于某个固定值（默认为1990年） | 去除无效数据，利于根据时间段的统计。 |

1. **标段信息基础表yw\_biaoduaninfo**

|  |  |  |  |
| --- | --- | --- | --- |
| **清洗项** | **清洗规则** | **目的** | **统计结果的影响** |
| 标段标识 | 必填 | 去除无效数据 | 对大数据统计无影响，可能对业务有影响。也可能在统计业绩时，无法做关联查询。 |
| 标段名称 | 必填 | 去除无效数据 | 对大数据统计无影响，可能对业务有影响。 |
| 开标日期 | 工程、政采的必填 | 去除无效数据 | 影响今日开标量等统计结果 |
| 不能晚于中标时间 | 去除非法数据 |
| 中标日期 | 如果存在中标金额，那就不能为空 | 利于统计某个时间段的中标情况。 | 对大数据统计无影响，可能对业务有影响。 |
| 晚于开标日期 | 去除非法数据 |
| 预算金额 | 工程、采购，大于中标金额 | 去除非法数据。 | 影响项目节约率、成交增值率的统计。 |
| 交易类，小于中标金额 | 去除非法数据。 |
| 中标金额 | 工程、采购，小于预算金额 | 去除非法数据、利于节约率的统计等 | 影响项目节约率、成交增值率的统计。 |
| 交易类，大于中标金额 | 去除非法数据 |
| 招标方式 | 符合“交易平台招标方式”、“交易平台采购方式”、“交易平台产权交易招标方式”、“交易平台土地交易招标方式”、“交易平台矿业招标方式”代码项 | 去除非法数据。 | 影响统计公开、邀请招标交易趋势 |
| 交易类型 | 符合“交易平台交易类型”代码项 | 去除非法数据。 | 影响交易类型的标段数统计 |

1. **招标公告基础表yw\_zhaobiaogginfo**

|  |  |  |  |
| --- | --- | --- | --- |
| **清洗项** | **清洗规则** | **目的** | **统计结果的影响** |
| 公告标题 | 必填 | 去除无效数据 | 对大数据统计无影响，可能对业务有影响。 |
| 公告发布时间 | 必填 | 去除无效数据 | 影响本月、本季度、本年的招标公告发布数及环比增长率的统计 |
| 公告截止时间 | 必填 | 去除无效数据 | 对大数据统计无影响，可能对业务有影响。 |

1. **变更公告基础表yw\_biangenggginfo**

|  |  |  |  |
| --- | --- | --- | --- |
| **清洗项** | **清洗规则** | **目的** | **统计结果的影响** |
| 公告标题 | 必填 | 去除无效数据 | 对大数据统计无影响，可能对业务有影响。 |
| 审核日期 | 必填 | 去除无效数据，利于增量抽取。 | 影响本月、本季度、本年的招标公告发布数及环比增长率的统计。影响大数据增量抽取 |
| 小于当前时间，大于某个固定值（默认为1990年） | 去除无效数据 |

1. **投诉信息基础表yw\_complaintinfo**

|  |  |  |  |
| --- | --- | --- | --- |
| **清洗项** | **清洗规则** | **目的** | **统计结果的影响** |
| 标段标识 | 必填 | 去除无效数据，可能影响到标段的关联 | 影响交易异常信息分析中的投诉信息的统计。 |
| 被投诉人 | 必填 | 去除无效数据，可能影响到被投诉单位的关联 | 对大数据统计无影响，可能对业务有影响。 |
| 受理人姓名 | 必填 | 去除无效数据 | 暂无影响 |
| 提交日期 | 必填 | 去除无效数据，利于统计某个时间段的投诉数量 | 暂无影响 |
| 小于审核时间，大于某个固定值（默认为1990年） | 去除无效数据，利于统计某个时间段的投诉数量 | 暂无影响 |
| 审核日期 | 必填 | 去除无效数据，利于增量抽取。 | 影响大数据增量抽取 |
| 小于当前时间，大于提交时间 | 去除无效数据，利于统计某个时间段的投诉数量 |

1. **场地预约表yw\_mtrprojectinfo**

|  |  |  |  |
| --- | --- | --- | --- |
| **清洗项** | **清洗规则** | **目的** | **统计结果的影响** |
| 预定标题 | 必填 | 去除无效数据 | 对大数据统计无影响，可能对业务有影响。 |
| 开标时间 | 必填 | 去除无效数据，利于类似于今日开标量之类的统计 | 对大数据统计无影响，可能对业务有影响。 |
| 小于评标时间，大于某个固定值（默认为1990年） | 去除无效数据，利于类似于今日开标量之类的统计 |
| 评标日期 | 必填 | 去除无效数据 | 对大数据统计无影响，可能对业务有影响。 |
| 大于开标时间 | 去除无效数据 |

1. **投标报名基础表yw\_toubiaodanwei**

|  |  |  |  |
| --- | --- | --- | --- |
| **清洗项** | **清洗规则** | **目的** | **统计结果的影响** |
| 标段标识 | 必填 | 去除无效数据 | 暂无影响 |
| 单位标识 | 必填 | 去除无效数据 | 暂无影响 |
| 报名时间 | 必填 | 去除无效数据 | 影响单体交易分析中的报名数量统计。也会导致按时间统计的报名数和累计报名数不同步。 |
| 小于缴纳保证金时间，大于某个固定值（默认为1990年） | 筛选出可能的异常数据，提高某个时间投标数的统计，避免总数和分时段统计数不对应的问题。 |
| 缴纳保证金时间 | 小于当前时间，大于报名时间 | 筛选出可能的异常数据 | 暂无影响 |

1. **中标公示基础表yw\_zhongbiaogsinfo**

|  |  |  |  |
| --- | --- | --- | --- |
| **清洗项** | **清洗规则** | **目的** | **统计结果的影响** |
| 公告标题 | 必填 | 去除无效数据 | 暂无异常数据 |
| 公告发布时间 | 必填 | 去除无效数据 | 影响本月、本季度、本年的中标公示发布数及环比增长率的统计 |
| 公告截止时间 | 必填，除了政采 | 去除无效数据 | 对大数据统计无影响，可能对业务有影响。 |
| 大于发布时间 | 去除无效数据 |
| 除了政采的，于发布时间的跨度不能超过半年。 | 筛选出可能的异常数据 |
| 审核日期 | 必填 | 去除无效数据，利于增量抽取。 | 影响本月、本季度、本年的中标公示发布数及环比增长率的统计。影响大数据增量抽取 |
| 小于当前时间，大于某个固定值（默认为1990年） | 去除无效数据 |

1. **语音通知信息表zj\_yytzpwinfo**

|  |  |  |  |
| --- | --- | --- | --- |
| **清洗项** | **清洗规则** | **目的** | **统计结果的影响** |
| 评委标识 | 必填 | 去除无效数据 | 对大数据统计无影响，可能对业务有影响。 |
| 抽取时间 | 必填 | 去除无效数据 | 暂无影响 |
| 小于应答时间或者开标时间，大于某个固定值（默认为1990年） | 筛选出可能的异常数据 | 暂无影响 |
| 应答时间 | 必填 | 去除无效数据 | 暂无影响 |
| 小于开标时间或在当天，大于抽取时间 | 筛选出可能异常的数据。 | 暂无影响 |
| 参加时间（即开标时间） | 大于抽取时间或在当天 | 筛选出可能的异常数据 | 暂无影响 |