**数据库设计**

主要包括案件表、人员表、单位表、代码数据字典表、组织机构表等

数据清洗时，需要注意以下说明：

1.单位代码说明：最高人民检察院 代码：100000，是最顶级单位。上下级关系根据pub\_organization中id和parent\_id进行关联。

2.数据字典表：pub\_code\_item，可以根据字段code\_type\_no查阅不同类别数据。

3.标红：字段名称红色必填。

4.批注：部分字段填充内容，进行了简单说明。

5.附件表：pub\_organization全国检察院机构表、pub\_code\_item数据字典表，全部已填充数据。

请严格按照上下说明进行数据清洗，否则会有问题特别是单位代码，如有问题，需要及时沟通，确保后续开发顺利。

表结构设计如下：

**xh\_caseinfo**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 字段名称 | 类型 | 描述 | OpenLaw | 提取 |
| UUID | varchar(64) | 案件id(主键) |  |  |
| AJXZ | varchar(64) | 案件性质（对应pub\_code\_item.ajxz） | E案由 | Excel与table取并集 |
| AJLX | tinyint | 案件类型(1单位性质案件 2个人性质案件) |  |  |
| UNITCODE | varchar(16) | 单位代（对应pub\_organization.code） | 单位名称  RV |  |
| TOP\_UNITCODE | varchar(16) | 省级单位代码（见pub\_organization表） | G法院 |  |
| HYLY | varchar(64) | 行业领域（行贿pub\_code\_item.hyly\_xh 受贿pub\_code\_item.hyly\_sh） | V庭审过程 |  |
| SAHJ | varchar(64) | 涉案环节（pub\_code\_item.sahj） | V庭审过程 |  |
| AFCS | varchar(64) | 案发场所（pub\_code\_item.afcs） | V庭审过程 | 使用table字段作为字典搜索 |
| LASJ | datetime | 立案时间 | V、T、R\* |  |
| PJSJ | datetime | 判决时间 | H判决日期 | 提取并转化为mysql的日期格式 |
| SADQ | varchar(64) | 涉案地区（pub\_code\_item. administrative\_area） | G法院\* | 匹配较为麻烦 |
| SASJ | datetime | 涉案时间 | V庭审过程\* |  |
| SAJE | decimal(12,2) | 涉案金额（万元） | V庭审过程\* |  |
| ZACS | decimal(10,0) | 作案次数 |  |  |
| DCFZ | tinyint | 多次犯罪 |  |  |
| SW | tinyint | 涉外（1是，2否） |  |  |
| AQZY | 1000 | 案情摘要 | V庭审过程\* | 前1000个字符 |
| CREATETIME | datetime | 创建时间 |  |  |
| UPDATETIME | datetime | 更新时间 |  |  |

**xh\_personinfo**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 字段名称 | 类型 | 描述 | OpenLaw | 提取 |
| UUID | varchar(64) | 人员id(主键) |  |  |
| AJID | varchar(64) | 案件id |  |  |
| DWID | varchar(64) | 单位id，有值说明此人为单位下的人员 |  |  |
| PERSONTYPE | tinyint | 1行贿人 2受贿人 | E案由 | 依据案由中的罪名判断 |
| XM | varchar(16) | 姓名 | J被告 |  |
| XB | varchar(64) | 性别(pub\_code\_item.gender) | R当事人 |  |
| CSRQ | datetime | 出生日期 | R当事人 | 注意日期格式 |
| AGE | tinyint | 年龄 | R当事人\* | 需要使用出生日期减去判决日期 |
| ZZMM | varchar(64) | 政治面貌（pub\_code\_item.gender） | R当事人 |  |
| RDDB | varchar(64) | 人大代表（全国、省级、地市、县级、乡镇、无） | R当事人 |  |
| ZXWY | varchar(64) | 政协委员（全国、省级、地市、县级、无） | R当事人 |  |
| RYSF | varchar(64) | 人员身份(pub\_code\_item.rysf) | R当事人 |  |
| ZW | varchar(128) | 职务 | R当事人 |  |
| ZJ | varchar(64) | 职级（见代码表） | R当事人\* |  |
| CARDTYPE | tinyint | 证件类型（1身份证，2其他） | R当事人 |  |
| CARDID | varchar(18) | 证件号码 | R当事人 |  |
| WHCD | varchar(64) | 文化程度（pub\_code\_item.whcd） | R当事人 |  |
| SZDQ | varchar(64) | 所在地区（pub\_code\_item. administrative\_area） | R当事人 |  |
| SSDW | varchar(128) | 所属单位（xxx公司） | R当事人 |  |
| XZZ | varchar(128) | 现住址 | R当事人 |  |
| SAJE | decimal(12,2) | 涉案金额（万元） | V庭审过程 |  |
| XSGX | varchar(64) | 与行受人关系（pub\_code\_item. xsgx） | V庭审过程 | 匹配 |
| CYSD | varchar(64) | 参与手段（pub\_code\_item. cysd） | V庭审过程 | 匹配 |

**xh\_unitinfo**

|  |  |  |
| --- | --- | --- |
| 字段名称 | 类型 | 描述 |
| UUID | varchar(64) | 单位id |
| AJID | varchar(64) | 案件id |
| DWMC | varchar(128) | 单位名称 |
| TYSHXYDM | varchar(20) | 统一社会信用代码 |
| FRXM | varchar(16) | 法人姓名 |
| FRSA | tinyint | 法人涉案（1是 2否） |
| UNITTYPE | tinyint | 1行贿方 2受贿方 |
| BCFJ | decimal(12,2) | 被处罚金（万元） |
| DWXZ | varchar(64) | 单位性质(pub\_code\_item. dwxz) |
| SZDQ | varchar(64) | 所在地区（pub\_code\_item. administrative\_area） |
| LICENSENO | varchar(20) | 营业执照号 |
| REGADD | varchar(128) | 注册地址 |
| WORKADD | varchar(128) | 办公地址 |