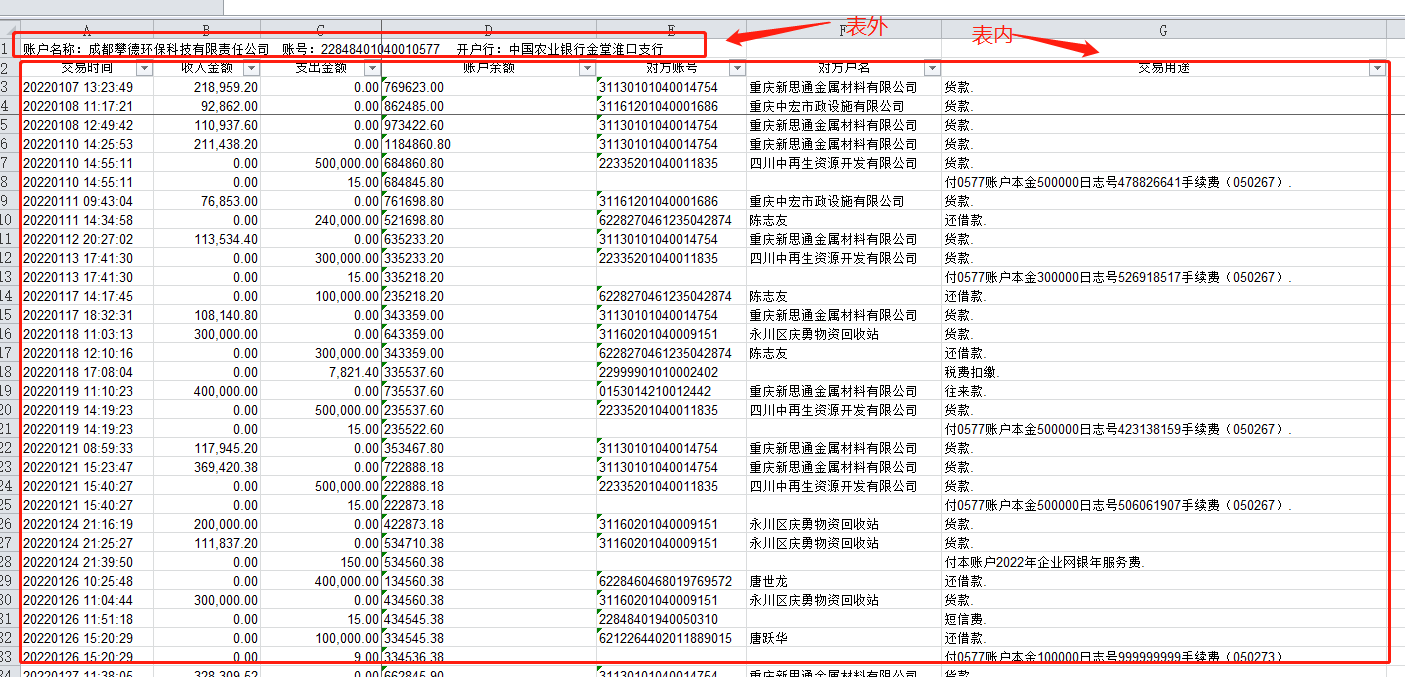
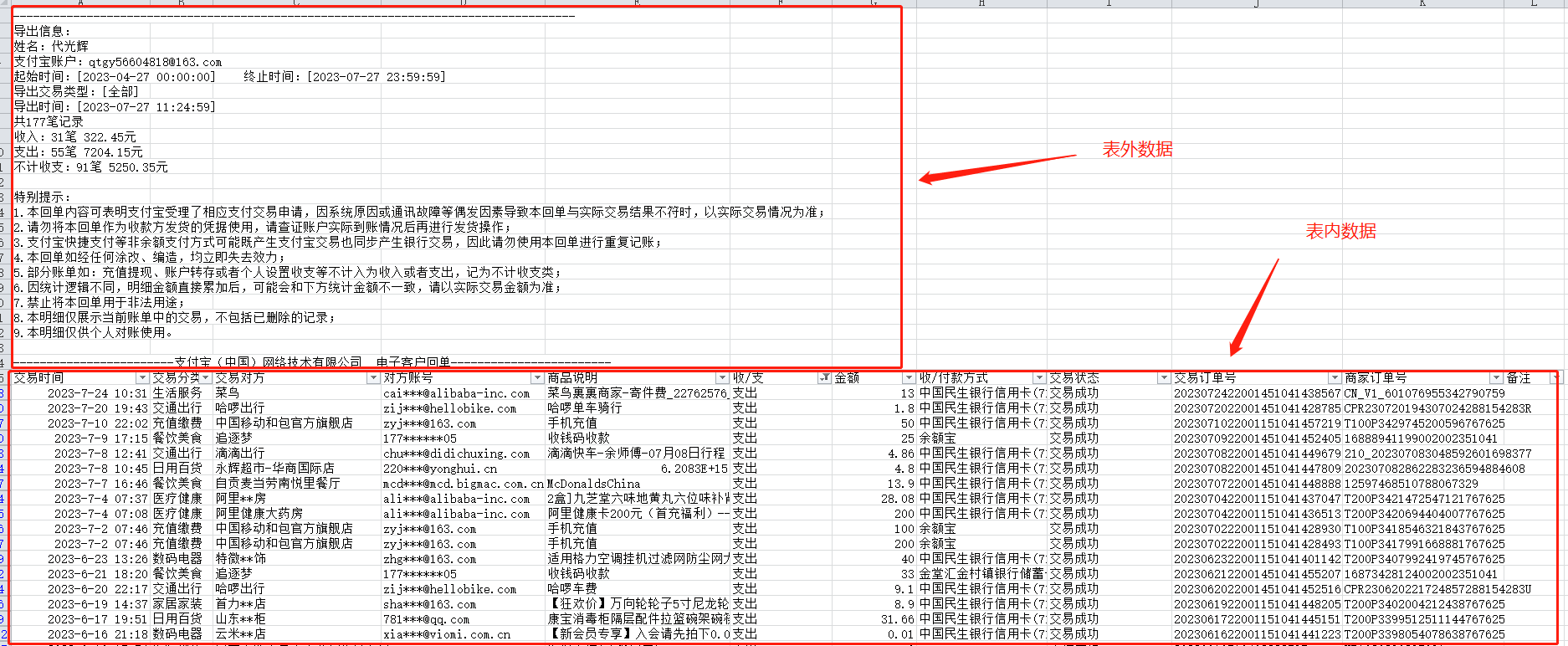
**需求：分析“银行流水数据”，从中筛选并提取指定信息**

**一. 银行流水数据样例：**





**二. 需要处理的数据：**

给到的数据data分为四部分，表外数据（res1）、表内数据（res2）和page\_sum、doc\_type四部分；

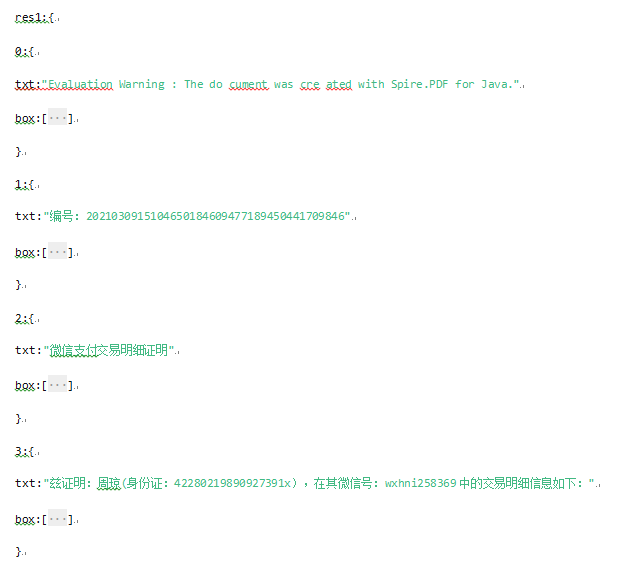


**1）res1数据：**

res1数据是指（表外数据），包含内容和坐标信息；（doc\_type值为image的是包含坐标信息的，excel和epdf的没有坐标信息；）0值是指res2前面的表外数据，1值是指res2后面的表外数据；

**2）res2数据：**

以下是res2数据（表内数据）的输出格式，包含每个单元格的行列信息和内容：

**** 

**3）page\_sum数据：**

显示一共识别处理的页数；

**4）doc\_type数据：**

doc\_type为文件类型，分为excel、epdf、image三种类型；不同类型体现在res1中是否包含坐标信息；

**三. 需求说明**

**需求1：从res1数据中，提取以下字段信息；**

输出格式参考第四部分的res1部分；

**1）agent\_type：bank，输出以下字段信息**

|  |  |  |
| --- | --- | --- |
| **字段** | **说明** | **输出内容** |
| agent\_type | 机构类型 | **bank：银行**  alipay：支付宝  wechat：微信支付 |
| bank\_name | 参考“金融机构信息对照表” | GY0001 |
| accout\_num | 银行账号号码 | 6228480329572520271 |
| account\_name | 账号的户名 | 李数 |
| date\_start | 起始日期 | 2022/05/22 |
| date\_end | 结束日期 | 2022/11/22 |
| account\_type | 账户类型 | 对私 |

**2）agent\_type：alipay的，输出以下字段信息**

|  |  |  |
| --- | --- | --- |
| **字段** | **说明** | **输出内容** |
| agent\_type | 机构类型 | bank：银行  alipay：支付宝  wechat：微信支付 |
| accout\_num | 支付宝账号号码 | caiyingbiao@hotmail.com |
| account\_name | 支付宝账号姓名 | 李数 |
| idcard\_num | 用户身份证号码 | 330724197805240011 |
| date\_start | 起始日期 | 2022/05/22 |
| date\_end | 结束日期 | 2022/11/22 |
| account\_type | 账户类型 | 对私 |

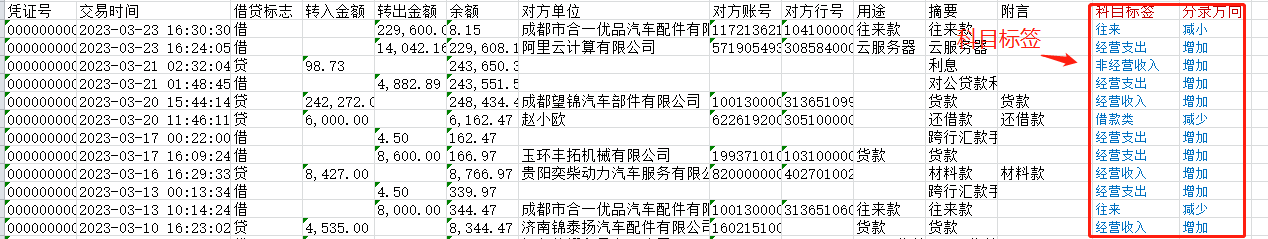
**3）agent\_type：wechat的，输出以下字段信息**

|  |  |  |
| --- | --- | --- |
| **字段** | **说明** | **输出内容** |
| agent\_type | 机构类型 | bank：银行  alipay：支付宝  wechat：微信支付 |
| accout\_num | 微信账号号码 | xshaitt |
| account\_name | 微信账号姓名 | 李数 |
| idcard\_num | 身份证号码 | 330724197805240011 |
| date\_start | 起始日期 | 2022/05/22 |
| date\_end | 结束日期 | 2022/11/22 |
| account\_type | 账户类型 | 对私 |

**需求2：分析res2数据，判断表头行，给非表头行数据，添加“科目标签”和“分录方向”；**

**res2输出；**

|  |  |  |  |
| --- | --- | --- | --- |
| **序号** | **字段名** | **说明** |  |
| 1 | header\_row | ”true”表示表头行，”false”表示非表头行； |  |
| 2 | accounting\_item | 科目标签，18个科目标签，查看附件《科目标签&居民消费支出分类》 |  |
| 3 | accouting\_entry | 分录方向 |  |
| 4 | trader\_nature | 交易对象的行业分类 |  |
|  |  |  |  |





**需求3：以规范格式输出res3；**

1）agent\_type：bank，

输出以下字段数据；（备注如果有多个，就输出多个）

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **交易日期** | **交易金额** | **收支类型** | **余额** | **备注（摘要）** | **对方账号** | **对方户名** |
|  | 2022/03/15 | 3571.42 | 收入 | 3571.42 | 跨行转账（文本） | 622841……8978（文本） | 张\*\*（文本） |
|  | 2022/03/16 | 100.00 | 支出 | 3471.42 | 支付宝（文本） | 2155\*\*690（文本） | 支付宝网络技术有限公司（文本） |

2）agent\_type：wechat，

输出以下字段数据（wechat的格式也不一定是不变的，根据原数据显示列，针对日期和金额格式做规范化处理）

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **交易日期** | **交易类型** | **收/支/其他** | **交易方式** | **交易金额** | **交易对方** |
|  | 2022/03/15（yyyy/mm/dd） | 商户消费（文本） | 支出（文本） | 中国银行（6633）（文本） | 3571.42（数值） | 王\*\* |
|  | 2022/03/16（yyyy/mm/dd） | 微信红包（文本） | 支出（文本） | 零钱（文本） | 3471.42（数值） | 发出群红包 |
|  | 2022/03/16（yyyy/mm/dd） | 二维码收款（文本） | 收入（文本） | / | 1500.00（数值） | 曾\*\* |

3）agent\_type：alipay，

输出以下字段数据（alipay的格式也不一定是不变的，根据原数据显示列，针对日期和金额格式做规范化处理）

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **收/支** | **交易对方** | **商品说明** | **收/付款方式** | **金额** | **交易日期** |
| 其他 | 天弘基金管理有限公司（文本） | 转账收款到余额宝（文本） | 余额宝（文本） | 1.77（数值） | 2022/03/16（yyyy/mm/dd） |
| 收入 | \*平（文本） | 转账（文本） | 余额宝（文本） | 500.77（数值） | 2022/03/16（yyyy/mm/dd） |
| 支出 | \*明（文本） | 收款（文本） | 中国建设银行(3955)（文本） | 1500.00（数值） | 2022/03/16（yyyy/mm/dd） |
| 支出 | 滴滴出行（文本） | 滴滴快车-王师傅（文本） | 花呗（文本） | 12.8（数值） | 2022/03/16（yyyy/mm/dd） |

**四、输出格式**

**1、一级参数**

|  |  |
| --- | --- |
| **参数名** | **说明** |
| success | true：代表执行成功  false：代表执行失败 |
| error-code | 返回错误码 |
| message | 报错信息 |
| data | 识别数据 |

**error-code说明**

|  |  |  |
| --- | --- | --- |
| **代码类型** | **信息** | **说明** |
| 401 | Type error | 文件格式不符合要求 |
| -1 | unknow | 未知错误 |

**2、data二级参数说明**

|  |  |  |
| --- | --- | --- |
| **参数名** | **说明** | **内容** |
| page\_sum | 总共页数 | 001 |
| doc\_type | 数据来源的文件格式，保存即可； | Image：图片  excel：电子表格  epdf：文本PDF文件 |
| res1 | 表外信息的数据提取 |  |
| res2 | 表内每行数据的以及科目标签 |  |
| res3 | 表内数据的指定字段信息提取 |  |

**3、三级参数：res1参数**

|  |  |  |
| --- | --- | --- |
| **字段** | **说明** | **输出内容** |
| agent\_type | 机构类型 | **bank：银行**  alipay：支付宝  wechat：微信支付 |
| bank\_name | 参考“金融机构信息对照表” | GY0001 |
| accout\_num | 银行账号号码 | 6228480329572520271 |
| account\_name | 账号的户名 | 李数 |
| date\_start | 起始日期 | 2022/05/22 |
| date\_end | 结束日期 | 2022/11/22 |
| account\_type | 账户类型 | 对私 |

|  |  |  |
| --- | --- | --- |
| **字段** | **说明** | **输出内容** |
| agent\_type | 机构类型 | bank：银行  alipay：支付宝  wechat：微信支付 |
| accout\_num | 支付宝账号号码 | caiyingbiao@hotmail.com |
| account\_name | 支付宝账号姓名 | 李数 |
| idcard\_num | 用户身份证号码 | 330724197805240011 |
| date\_start | 起始日期 | 2022/05/22 |
| date\_end | 结束日期 | 2022/11/22 |
| account\_type | 账户类型 | 对私 |

|  |  |  |
| --- | --- | --- |
| **字段** | **说明** | **输出内容** |
| agent\_type | 机构类型 | bank：银行  alipay：支付宝  wechat：微信支付 |
| accout\_num | 微信账号号码 | xshaitt |
| account\_name | 微信账号姓名 | 李数 |
| idcard\_num | 身份证号码 | 330724197805240011 |
| date\_start | 起始日期 | 2022/05/22 |
| date\_end | 结束日期 | 2022/11/22 |
| account\_type | 账户类型 | 对私 |

**4、res2参数**

|  |  |  |  |
| --- | --- | --- | --- |
| **序号** | **字段名** | **说明** | **内容** |
| 1 | row\_order | 当前行的位置，页码+行序 | 003\_8 |
| 2 | 1.1  1.2  2.1 | 当前单元格的位置，行序+列序 | 1.1  1.2  2.1 |
| 3 | account\_label | 科目标签 | 详情查看如下 |
| 4 | row\_label | 行标签 | 详情查看如下 |
|  |  |  |  |

**1）row\_label的详细参数**

|  |  |  |  |
| --- | --- | --- | --- |
| **序号** | **字段名** | **说明** | **内容** |
| 1 | header\_row | 是否是表头行，true代表是，false代表不是 | true  false |
| 2 | trader\_nature | 交易对象的行业分类 | 快餐服务\眼镜制造 |
| **3** | **in\_spend\_type** | **收支类型** | **收入**  **支出** |

**2）account\_label的详细参数**

|  |  |  |  |
| --- | --- | --- | --- |
| **序号** | **字段名** | **说明** | **内容** |
| 1 | N-accounting\_item | 主科目标签 |  |
| 2 | N-accounting\_entry | 主分录方向 |  |
| 3 | accounting\_item | 副科目标签 | 储蓄类资产 |
| 4 | accounting\_entry | 副分录方向 | 增加 |
|  |  |  |  |

**5、res3参数**

|  |  |  |  |
| --- | --- | --- | --- |
| **序号** | **字段名** | **说明** | **内容** |
| 1 | row\_order | 当前行的位置，页码+行序 | 003\_8 |
| 2 | 1.1  1.2  2.1 | 当前单元格的位置，行序+列序 | 1.1  1.2  2.1 |
| 3 | header\_row | 是否是表头行，true代表是，false代表不是 | true  false |
|  |  |  |  |

**五、交付内容**

1、源代码

2、镜像文件

3、部署文档（源码、镜像的部署文档&接口文档）