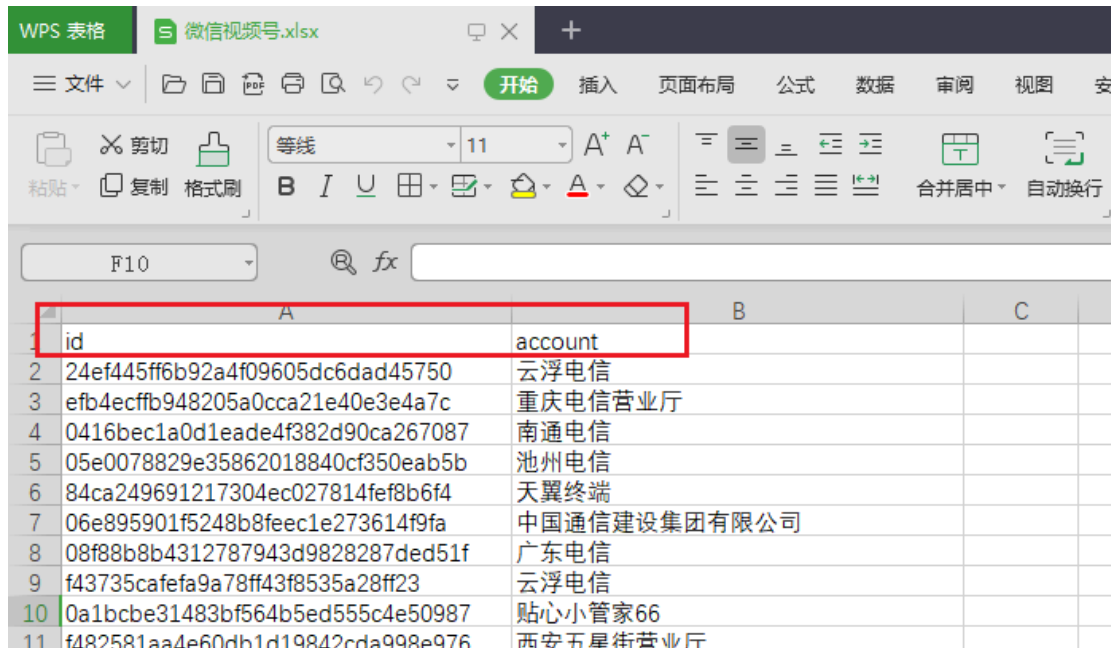


微信视频号采集流程

1. 提供账号模板

- 按照如下格式，提供要采集的微信视频号账号列表



id	account
24ef445ff6b92a4f09605dc6dad45750	云浮电信
efb4ecffb948205a0cca21e40e3e4a7c	重庆电信营业厅
0416bec1a0d1eade4f382d90ca267087	南通电信
05e0078829e35862018840cf350eab5b	池州电信
84ca249691217304ec027814fef8b6f4	天翼终端
06e895901f5248b8feec1e273614f9fa	中国通信建设集团有限公司
08f88b8b4312787943d9828287ded51f	广东电信
f43735cafefa9a78ff43f8535a28ff23	云浮电信
0a1bcbe31483bf564b5ed555c4e50987	贴心小管家66
f482581aa4e60dh1d19842rda998e976	西安五星街营业厅

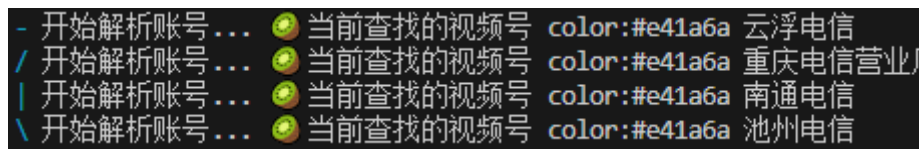
- 表行首行 `id` 和 `account` 两列，这两列的值不要有空格
- 命名为 `微信视频号.xlsx`

2. 安装依赖

```
# 执行npm install或者cnpm install安装依赖
npm install
```

3. 放入账号模板解析账号

- 将 `微信视频号.xlsx` 放入 `read_data` 目录下，替换原有xlsx
- 执行 `npm run read-excel`，读取excel，读取成功后把会excel中的数据，生成到当前目录的 `account.json` 中，自行验证



```
- 开始解析账号... 当前查找的视频号 color:#e41a6a 云浮电信
/ 开始解析账号... 当前查找的视频号 color:#e41a6a 重庆电信营业厅
| 开始解析账号... 当前查找的视频号 color:#e41a6a 南通电信
\ 开始解析账号... 当前查找的视频号 color:#e41a6a 池州电信
```

- 终端出现 `解析写入account.json成功!` 代表账号解析完成

4. 设定采集时间

- 查看 `config.js`
- 修改 `num` 的值，默认为7，表示采集近7天的数据

5. 执行采集

- 运行 `npm run start-collect`
- 等待采集完成

6. 查看采集结果日志

- 执行 `npm run count-analysis`

```
[2024-07-02T16:54:14.456] [INFO] default - excel中视频号数量: 367 个
[2024-07-02T16:54:14.457] [INFO] default - excel中视频号数量去重之后: 210 个
[2024-07-02T16:54:14.458] [INFO] default - 可以正常读取或搜索到的账号: 198 个
[2024-07-02T16:54:14.460] [INFO] default - 不能正常读取(视频号不存在)的账号: 12个
[2024-07-02T16:54:14.460] [INFO] default - 已读取, 成功采集的账号: 4个,
[2024-07-02T16:54:14.461] [INFO] default - 已读取, 没有进行采集的账号(可能是日期内无发布, 可能是其他原因)的账号: 194个,
[2024-07-02T16:54:15.476] [INFO] default - 本次采集的所有excel合计总条数 0
```

7. 查看结果文件

- `output_data` 中存放的就是采集的数据文件