

csv2json模块实战使用

数字世界，一切都是数据，一切都是对象；json对象，就像开启数字世界的钥匙。

csvtojson模块的强大功能，在回车的瞬间就被其效率惊叹到，再查看其结果更是让你满血沸腾，2500多行就是秒间，一眨眼的功夫完成，统一齐整的json格式。早之前，在使用python pandas功能时，曾经为其强大和灵活的功能所震过，记忆犹新；因此在使用“csvtojson”前，还想计划先用pandas将其转换为json对象，保存文档交nodejs开发使用，当看到结果后，完全不需要了，已经完完全全超过想象的效果，甚至更强。在此推荐给小白们使用，特别时遇到有大量重复的，有统一数据格式的属性的，类似excel和csv文档数据需要程序处理的，这将大大提高效率，感受一下计算机带给我们的“秒”感。以下是本人使用的过程和心得，分享给大家。

第一步：npm install csvtojson，安装成功后，同时引入“fs”模块，定义好“源csv文件路径”和要转换的目标“json对象文件路径”；

一条指令（仅一条指令！），csv().from(path="csv源文件").then(=>Promise)，按Promise格式对json对象文件写入到目标路径（path="json对象文件"）；node执行后，命令行中显示“文件已保存”，即：转换成功。

```
const csvFilePath = './weixin_data/jlwx0612.csv';
const jsonFilePath = './weixin_data/jielong_wxin.json'; // 指定保存 JSON 的文件路径
const fs = require('fs');
const csv = require('csvtojson');

/**
 * csv()
 * .fromFile(csvFilePath)
 * .then((jsonObj) => {
 *   // 将 JSON 对象写入文件
 *   fs.writeFile(jsonFilePath, JSON.stringify(jsonObj), (error) => {
 *     if (error) {
 *       console.log(error);
 *     } else {
 *       console.log('文件已保存');
 *     }
 *   });
 * });
 */
```

第二步：按json对象目标文件保存的路径，点击文件，其结果显示为一数组（LIST）文件，数组中每个元素均为相同格式的json对象。

- csv 源文件格式：
- 源文件中有四个属性，即Excel或csv文件里的4个列，分别为"full_name","memo","context"和"duration"；每一行有四个用逗号隔开的字符。

- | | full_name | memo | context | duration |
|---|---------------|-------------------|---------------------|----------|
| 1 | WMM,2023/6/12 | ,整理共学共创贡献度 | ,增录 Docusaurus 视频一篇 | 3h,3 |
| 3 | KGZ,2023/6/12 | ,python基础第1,2章及作业 | 2h",2 | |
| 4 | AMW,2023/6/12 | ,docusaurus 站点 | 2.5h,2.5 | |
| 5 | KLZ,2023/6/12 | ,用中学,实际上手ios 开发 | 180分钟 | 3 |
| 6 | AMD,2023/6/12 | ,页面结构 | 1h,1 | |
| 7 | LJZ,2023/6/12 | ,docusaurus | 1小时 | 1 |

- 转换后的json对象文件格式：

- 转换后，在csv里每一行的记录，按照对应属性，组成4个键值对（key-value）即：csv文件的4个列（属性）分别对应4个key；每个key-value对，value为每行里对应的值,通过csv文件中逗号","区隔分别赋值，如所示：

```
[
  {"full_name": "WMM", "memo": "2023/6/12", "context": "整理共学共创贡献度，增录 Docusaurus 视频一篇3h", "duration": "3"},
  {"full_name": "KGZ", "memo": "2023/6/12", "context": "python基础第1,2章及作业2h", "duration": "2"},
  {"full_name": "AMW", "memo": "2023/6/12", "context": "docusaurus 站点 2.5h", "duration": "2.5"},
  {"full_name": "KLZ", "memo": "2023/6/12", "context": "用中学，实际上手iOS 开发，180分钟", "duration": "3"},
  {"full_name": "AMD", "memo": "2023/6/12", "context": "页面结构 1h", "duration": "1"},
  {"full_name": "LJZ", "memo": "2023/6/12", "context": "docusaurus 1小时", "duration": "1"}
]
```

- 至此，我们就可以灵活使用json对象文件，按照属性选择，做下一步的开发使用；如:统计接龙时长，统计每天的接龙人次等。同时通过full_name属性，关联对应的user_id,"identity_number"等唯一标识，补充完整接龙记录。为后续的数据清理和运营创造了条件。