

# DataFrame操作API

---

## DataFrame管理

### 1. 获取DataFrame列表

- **端点:** `/api/dataframes/list`
- **方法:** GET
- **描述:** 获取所有可用的DataFrame变量名列表
- **响应示例:**

```
{
  "status": "success",
  "data": [
    "df1",
    "df2",
    "df3"
  ],
  "message": "获取DataFrame列表成功"
}
```

### 2. 获取DataFrame信息

- **端点:** `/api/dataframes/info/{name}`
- **方法:** GET
- **参数:**
  - `name`: DataFrame变量名
- **描述:** 获取指定DataFrame的详细信息
- **响应示例:**

```
{
  "status": "success",
  "data": {
    "basic_info": {
      "行数": 1000,
      "列数": 10,
      "内存占用": "1.2 MB"
    },
    "columns": [
      {
        "name": "column1",
        "type": "int64",
        "null_count": 0
      }
    ],
    "preview": {
      "head": [
```

```
{
  {
    "column1": 1,
    "column2": "value"
  },
  "summary": {
    "column1": {
      "count": 1000,
      "mean": 50.5,
      "std": 28.86,
      "min": 1,
      "25%": 25.75,
      "50%": 50.5,
      "75%": 75.25,
      "max": 100
    }
  }
},
"message": "获取信息成功"
}
```

### 3. 获取DataFrame预览

- **端点:** `/api/dataframes/preview/{name}`
- **方法:** GET
- **参数:**
  - `name`: DataFrame变量名
- **描述:** 获取DataFrame的预览信息
- **响应示例:**

```
{
  "status": "success",
  "data": {
    "shape": [1000, 10],
    "columns": {
      "column1": "int64",
      "column2": "object"
    },
    "memory_usage": 80000,
    "sample_data": {
      "column1": [1, 2, 3, 4, 5],
      "column2": ["a", "b", "c", "d", "e"]
    }
  },
  "message": "获取预览信息成功"
}
```

### 4. 保存DataFrame

- **端点:** `/api/dataframes/{name}/save`
- **方法:** POST
- **参数:**
  - `name`: DataFrame变量名
- **描述:** 将DataFrame保存为文件
- **请求体:**

```
{
  "file_path": "output.csv",
  "file_type": "csv",
  "save_options": {
    "index": false,
    "encoding": "utf-8"
  }
}
```

- **响应示例:**

```
{
  "status": "success",
  "data": {
    "file_path": "output.csv",
    "size": 1024
  },
  "message": "DataFrame保存成功"
}
```

## 数据类型处理

DataFrame API支持以下数据类型的自动处理:

### 1. 基本类型:

- 整数 (int8, int16, int32, int64)
- 浮点数 (float32, float64)
- 布尔值 (bool)
- 字符串 (object)

### 2. 特殊类型:

- 时间戳 (datetime64[ns])
- 类别 (category)
- 空值 (None/NaN)

### 3. 数组类型:

- NumPy数组
- 列表

所有返回的数据都会被自动转换为JSON安全的格式。

## 错误处理

所有API都可能返回以下错误：

- 400: 请求参数错误
- 404: DataFrame不存在
- 500: 服务器内部错误

错误响应格式：

```
{
  "status": "error",
  "data": {},
  "message": "错误描述"
}
```

## 注意事项

1. DataFrame操作是内存中的操作，注意内存使用
2. 大数据集建议使用分批处理
3. 保存文件时注意磁盘空间
4. 预览功能默认只返回前5行数据
5. 统计信息会自动处理空值和异常值