

CSV 智能可视化分析系统 PRD (产品需求文档)

1. 产品概述

CSV 智能可视化分析系统是一款基于 FastAPI 后端与 Streamlit 前端实现的数据分析与可视化工具。用户可上传 CSV 文件，系统自动进行统计分析、图表生成，并调用 GPT 模型生成中文数据分析总结报告。

核心功能包括：

- CSV 数据解析与统计
- 自动生成直方图、柱状图
- 基于 GPT 的智能中文数据分析
- 可视化前端展示，便于快速查看数据与分析结果

2. 产品目标

- 提供企业或个人用户一个轻量级的 CSV 数据智能分析工具
- 自动生成统计信息、图表和分析报告，减少手动分析工作
- 支持在线可视化展示，增强数据洞察力
- 前端操作简洁、交互友好

3. 用户角色 & 使用场景

角色

- **数据分析师 / 业务人员**：快速查看 CSV 数据分布、生成可视化报告
- **开发者 / 产品人员**：搭建内部数据可视化工具或演示项目
- **教育与研究用户**：学习数据分析与可视化方法

场景

- CSV 数据快速统计与可视化分析

- 教学或演示数据分析流程
 - 对大批量 CSV 文件进行批量分析和对比
-

4. 产品功能描述

4.1 CSV 上传与解析

- 支持 CSV 文件上传
- 对上传文件进行格式检查
- 解析 CSV 数据为 Pandas DataFrame
- 提取基本信息：
 - 行数、列名
 - 每列非空数量
 - 数值列描述性统计
 - 分类别唯一值数量
- 上传失败时，返回错误提示

4.2 数据统计与图表生成

- **数值列**
 - 自动生成直方图
 - 支持自适应分箱（30 个）
 - 显示频数和分布形态
- **分类列**
 - 自动生成柱状图（限制显示前 15 类）
 - 标签旋转优化可读性
- 图表统一设置大小和布局，保证美观

4.3 GPT 智能分析报告

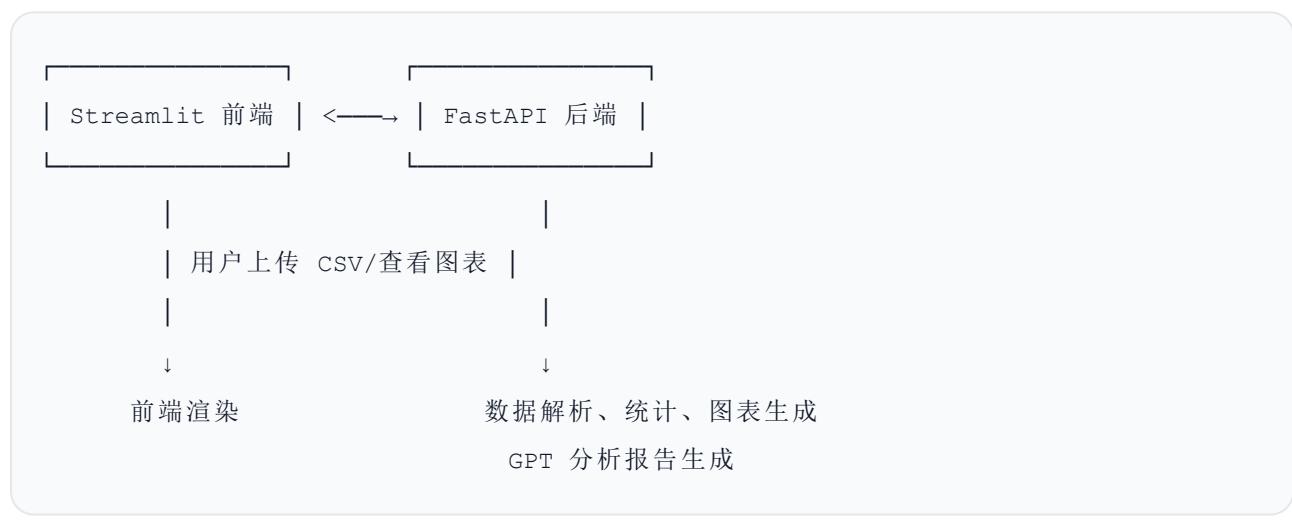
- 自动调用 GPT 模型生成中文分析总结
-

- 支持快速辅助业务分析与决策

4.4 前端可视化展示（Streamlit）

- 上传 CSV 文件入口
- 数据预览表格
- 统计信息折叠面板
- 图表展示区
- GPT 中文分析总结
- 响应式布局，支持宽屏显示

5. 系统架构



6. 技术栈

后端

- FastAPI
- Pandas
- Matplotlib
- OpenAI Qwen 模型

前端

- Streamlit
 - requests
 - Base64 图表展示
-

7. 性能指标

指标	目标
CSV 文件解析速度	10000 行 < 5 秒
图表生成延迟	< 500ms
单次上传文件数	≥ 5 个
GPT 分析生成速度	< 5 秒

8. 后续规划

- 增加 Excel 文件支持
 - 增加多图表组合分析
 - 增加数据筛选与过滤功能
 - 前端交互优化与主题支持
-