

# 个人近一年消费数据分析报告

## 项目概述：

- 关键词：Python · Pandas · 数据清洗 · 可视化 · 探索性分析
- 项目周期：2026 年 2 月 · 独立完成 · 总计约 20 小时
- 价值：通过分析近一年、3128 条（实际有效条数 2855 条）消费记录，识别出 16%的月度支出优化空间，建立个人财务管理模型，实现合理消费

## 分析内容与目标：

- 消费情况：钱主要花在哪里？何时容易过度，冲动消费？哪些支出可优化？
- 分析任务：支出结构分析、时间模式挖掘、可优化点识别
- 产出标准：形成可视化报告，提出具体、可执行的优化建议

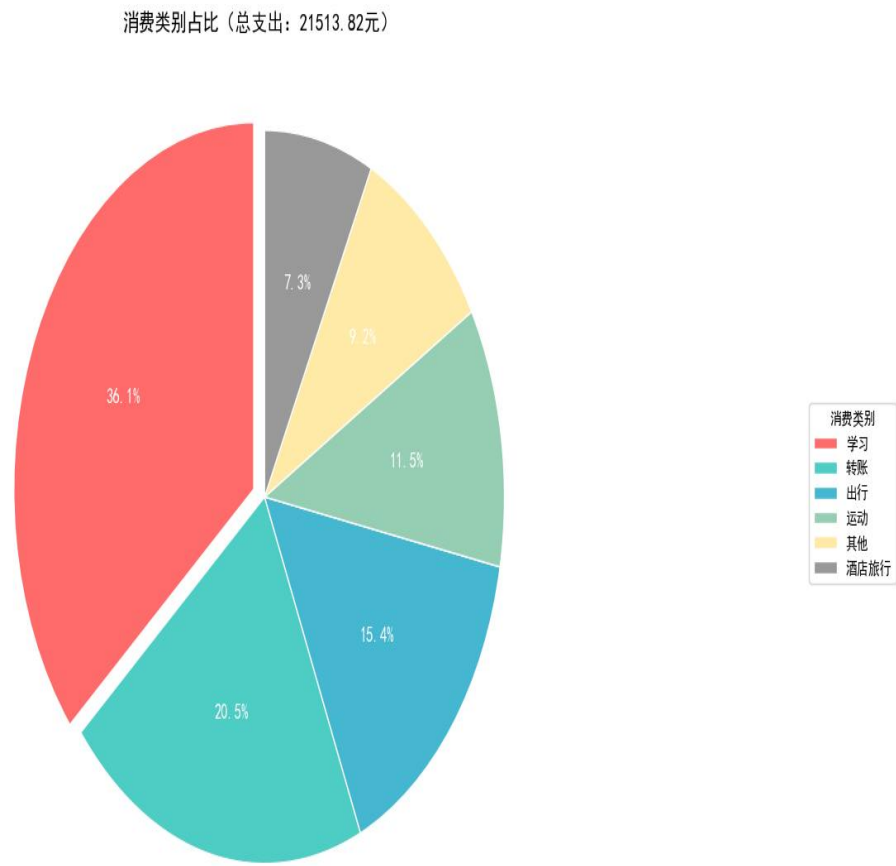
## 分析过程：

- 1.数据准备：导出支付宝近一年账单，用 `pd.read_csv` 加载数据至 pandas
- 2.输出前 5 行，了解原始数据形状，缺失值情况，列名等基本信息
- 3.数据清洗：删除无用列 'Unnamed: 8'，  
筛选“收支类型 = 支出”的数据，  
清洗「金额」字段（去除非数字字符），过滤金额 $\leq 0$ ，剩余数据量：2855 行，为有效数据  
转换「记录时间」为日期格式  
统一「分类」名称，分为学习，转账，出行，运动，旅行，其他六类，

进行消费结构分析

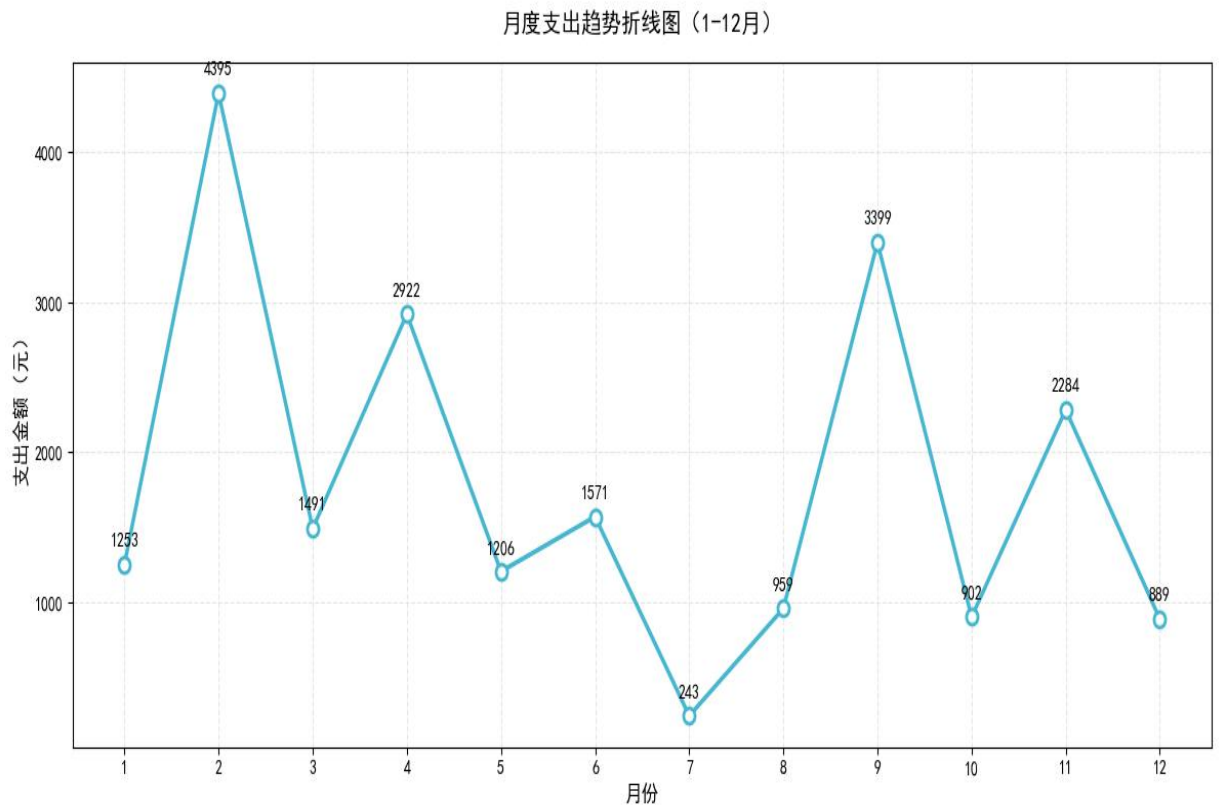
分析结果：

1.消费结构



结论与建议：转账类消费占比达 20%，多为朋友聚餐后 AA，以后可以适当降低聚餐频次

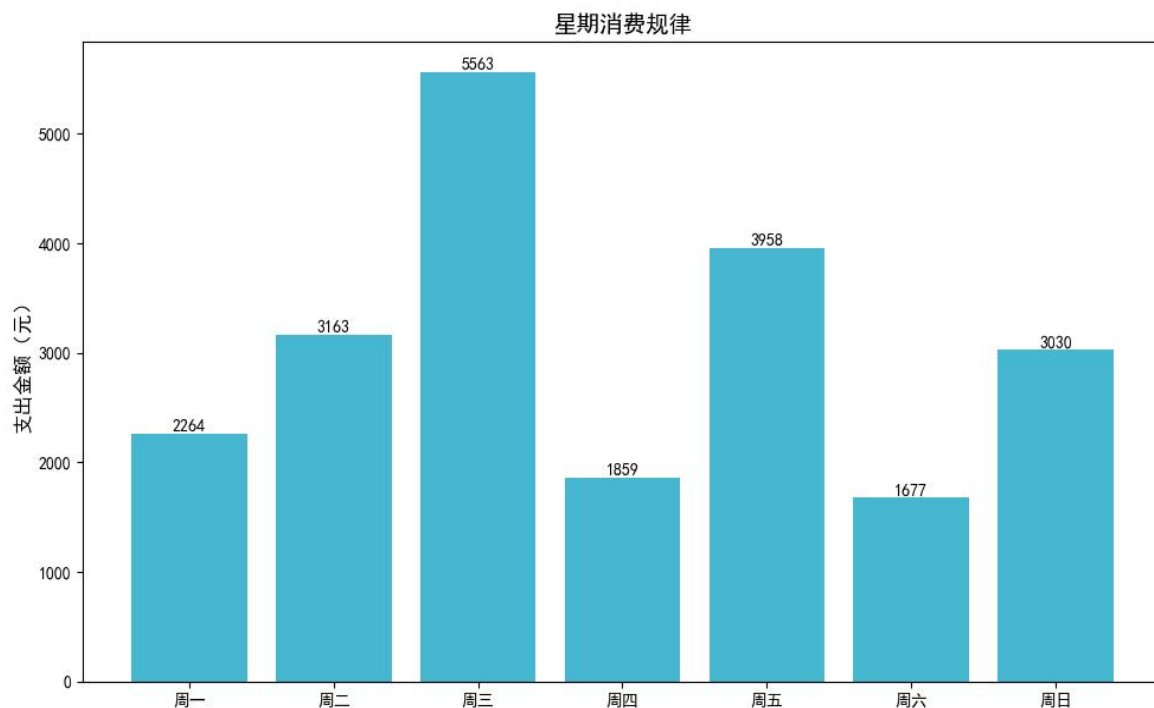
2.月度消费趋势



结论与建议：2月与9月均为假期与学期过渡期，朋友聚餐多，4月和9月有假期旅游，导致支出偏高

3.消费集中度：金额最大的前10笔支出中有7笔属于旅游类，且均为飞机票，以后可以多采用火车出行

4.星期消费趋势



结论与建议：因为周三，周五课少，聚餐集中于这两天

## 最终结论与行动

### 结论

- 可优化空间：每月约 **16%**的支出（主要来自聚餐与出行方式）
- 行动建议：设立聚餐预算上限、建立周三，周五消费预警

### 项目复盘

- 技术收获：掌握了数据导入，**Pandas** 数据清洗、统一消费分类为学习，转账，出行等 **6** 类、时间序列分析、数据可视化，总结分析与建议，自己完成了数据分析完整流程
- 难点突破：选择什么图表来进行可视化，如何提高消费自动分类边界的准确率
- 可改进点：未来可加入收入数据计算储蓄率，或建立自动化监控系统