

第一阶段作业

1. 任务要求

针对拉勾网2022-2023年的数据集做可视化分析。要求覆盖：

- 数据概况（取值分布、特征数量、缺失值、样本数）
- 单变量分析（按照类别和数值型区分开分析，其中部分内容需要做预处理后才可以分析，如薪资）
- 多变量分析（抽取几个变量进行对照分析）

要求本阶段任务必须完成5~6个有价值的问题的挖掘，主题自拟，以下是例子：

1. 研究生和本科生在薪资方面的差距真的很大吗？（不同的工作岗位对于不同学历下的薪资差距分布是怎样的呢？）
2. 大厂和小企业在薪资待遇上真的差距很大吗？
3. 薪资和哪些因素密切相关？
4. 哪些工作岗位需求量比较大？
5.

2. 输出要求

- 以个人为单位输出，不要雷同。分析问题的多少、程序复杂度、分析问题的深度、可视化的新颖性评分。
- 提交.ipynb文件，分析结论使用Markdown格式书写在代码附近。每个问题要使用醒目标题（##）
- 文件名为：学号_姓名（作业1）.ipynb