## A组大数据电商应用实训总结

**一、项目概述**

1在数字经济时代，电商行业已成为全球经济增长的核心引擎，2023年中国网络零售额突破15万亿元，占社会消费品零售总额的27.6%，凸显其重要地位。数据作为电商运营的"中枢神经"，通过用户行为分析、库存优化和精准营销驱动业务增长。实训大数据大屏应运而生，实时可视化交易、流量及用户画像等关键指标，帮助运营者快速决策，提升转化率与用户体验，是数字化运营的重要工具。

2. 周世攀负责获取京东的商品数据，数据清洗部分由覃思倩担任

卢官福处理数据分析，数据可视化由陆甜甜设计，Web数据大屏开发由谢明焕完成

**二、数据获取清洗**

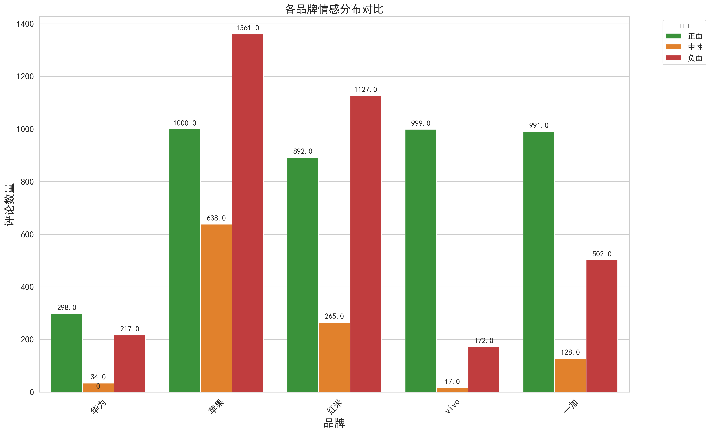
1.小组采用开源库Drissionpage完成传统selenium与request爬虫的操作，采用pandas进行数据清洗

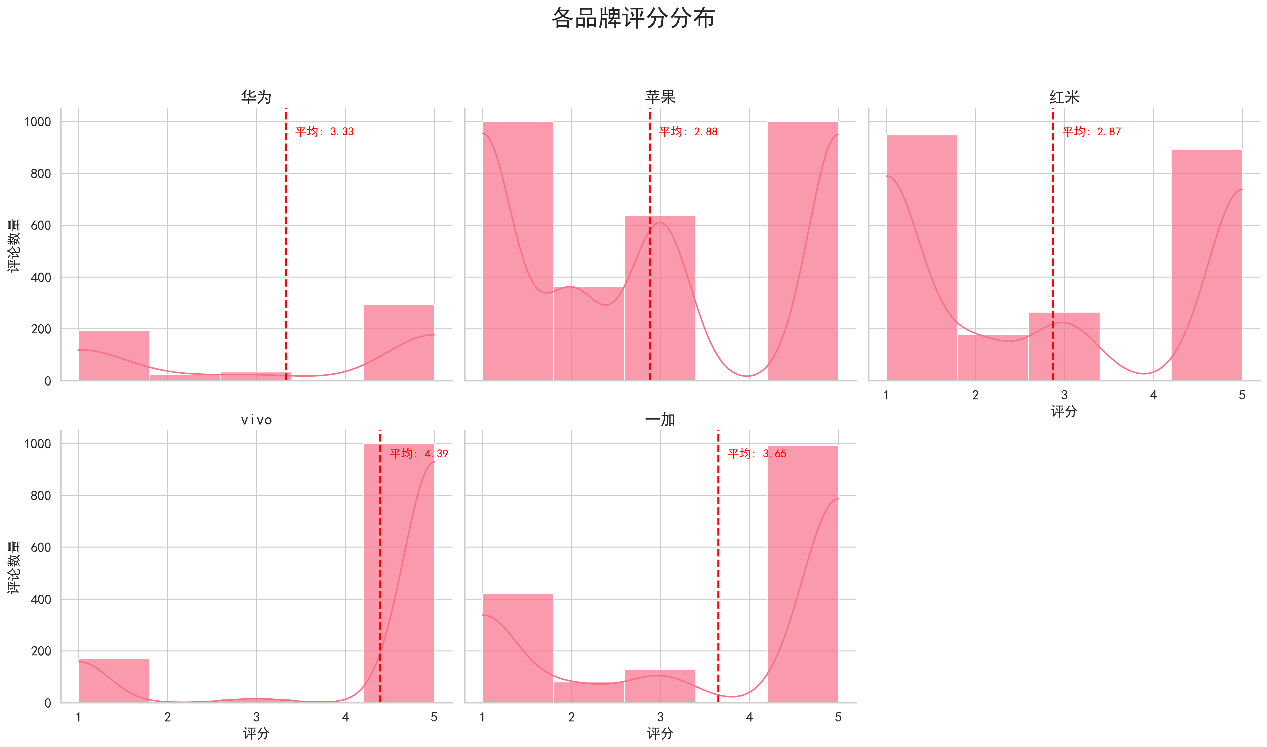
2. 数据由两部分成分，评分，以及评论内容，数据记录条数为8692条，数据类型为用户评价文本，评论内容能有效挖掘用户所倾向的内容，可帮助企业寻找自身缺点并解决，也能使企业放大自身品牌优势

**三、数据分析方法**

1.介绍小组所采用的数据分析方法，并解释选择这些方法的原因及其适用性。

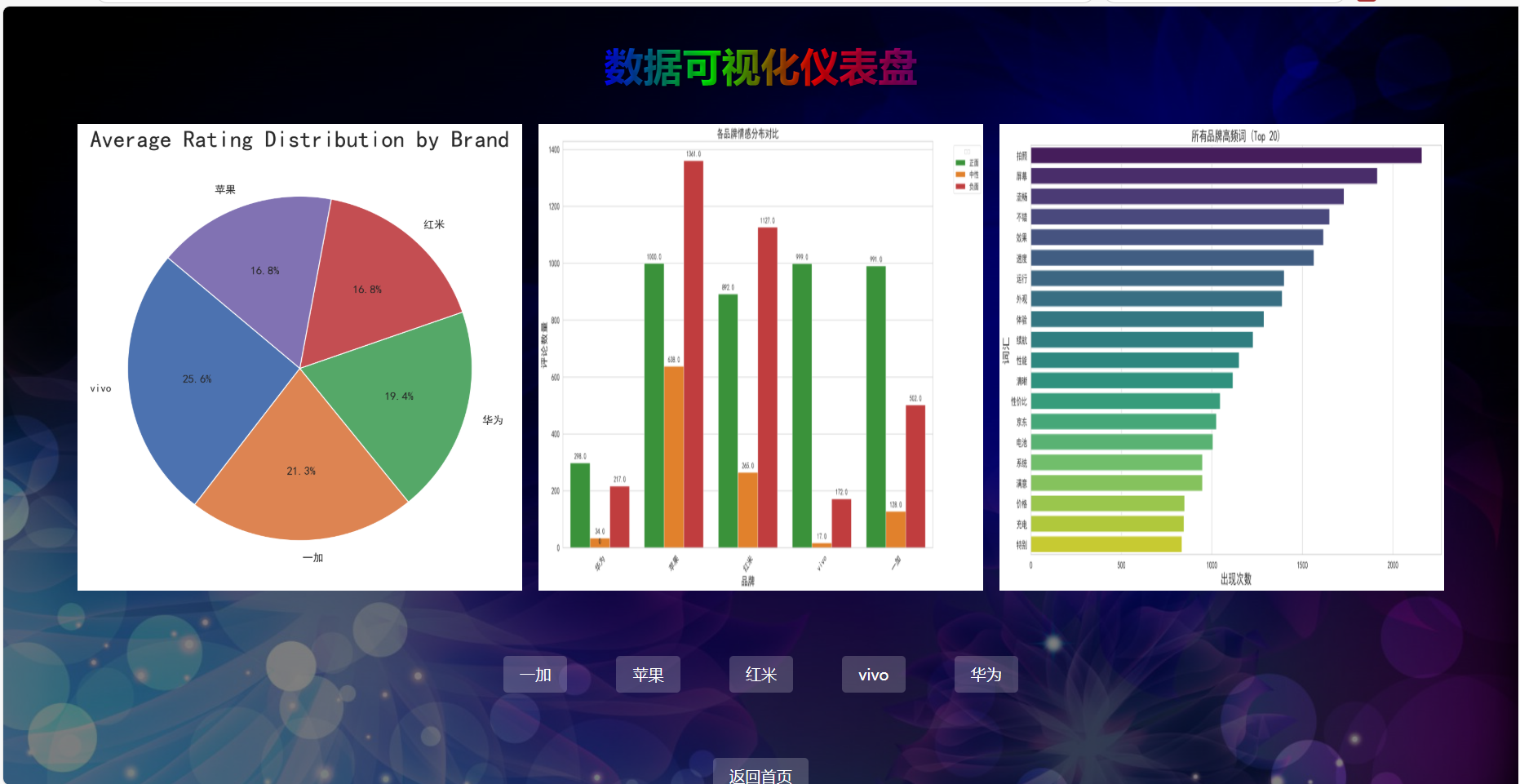
采用了监督学习法，将评论分为正面、中性、负面三类（基于评分标签吗，并使用朴素贝叶斯（MultinomialNB）作为分类模型





四、数据大屏可视化实现

1. 采用响应式布局，以半透明黑色背景增强数据可视化的对比度，搭配渐变色标题提升视觉吸引力。通过Flexbox实现图片和按钮的灵活排列，确保多终端适配。顶部展示核心指标图表（评分分布、情感分析、关键词），下方设置品牌导航快捷入口，整体风格简洁现代，聚焦数据呈现。



**五、团队协作与沟通**

在最开始的数据爬取部分，一开始只能通过复制粘贴代码块部分才能爬取数据，但是多次爬取还有被暂时禁止访问的风险，后续经过学习与ai的帮助下才完成了对爬虫的封装。

**六、成果与收获**

此次实训让我与同学之间的友谊更加深厚，并学习了课外的知识，体验了在实际企业中的工作流程。在实训过程中，施宇老师也教会了我们许多，如何更好地使用ai，如何寻找数据集，如何部署数据可视化大屏