《电商销售趋势分析代码测试报告》

一、测试概述

本次测试针对电商销售趋势分析代码，重点验证其在分布式环境下对电商数据的处理、分析、可视化及结果保存功能。

二、测试环境

2.1 硬件

分布式集群节点：CPU [具体型号]，核心数 [3]，主频 [X] GHz；内存 [60] GB；存储 [X] TB。

客户端机器：CPU [英特尔i5]，核心数 [X]，主频 [X] GHz；内存 [X] GB；存储 [X] GB。

2.2 软件

操作系统：分布式集群为 [具体 Linux 发行版及版本]，客户端为 [具体操作系统及版本]。

Java、Python、PySpark、Matplotlib、Seaborn、Pandas 均为 [对应版本]。

Hadoop 分布式文件系统版本为 [具体版本]。

三、测试用例及结果

测试模块 测试用例 预期结果 实际结果 是否通过

数据读取与预处理 正常数据读取 成功读取 E\_commerce.csv 文件，数据框含预期列 [是]

日期列处理 Order Date 转换为日期类型，正确提取 Year 和 Month [是]

异常值处理 Discount 列异常值 'xxx' 转换为 0.0 [是]

空值和非数值处理 Quantity 列空值和非数值转换为 0 [是]

各维度销售数据分析 总体销售趋势分析 按年和月分组计算总销售额、平均销售额和订单数量，结果数据框格式正确 [是]

按产品类别的销售分析 按产品类别和产品分组计算总销售额和订单数量，显示最畅销产品，结果数据框格式正确 [是]

其他维度分析（按地区、客户行为等） 对应分组计算相应指标，结果数据框格式正确[是]

数据可视化 总体销售趋势可视化 弹出月度销售趋势折线图，标题、坐标轴标签和数据显示正确[是]

其他可视化（折扣与销售关系等） 对应图表弹出，标题、坐标轴标签和数据显示正确 [是]

分析结果保存 结果保存到分布式文件系统 在指定路径生成分析结果 CSV 文件 [是]

四、测试总结

4.1 通过率

本次测试共 [9] 个用例，通过 [9] 个，通过率为 [100%]。

4.2 未通过用例分析

未通过用例编号 问题描述 可能原因 解决方案建议

[编号] [具体问题] [可能原因] [解决方案]

五、结论与建议

5.1 结论

代码大部分功能正常，但存在部分未通过用例，需修复优化。

5.2 建议

修复未通过用例的问题，确保代码准确性和稳定性。

增加异常处理机制，提高代码健壮性。

优化代码性能，处理大规模数据时采用高效算法和数据结构。

编写详细文档，方便后续维护扩展。