题目一：电商订单数据清洗与多维度分析

题目描述：

现有一份电商订单数据（orders.csv），包含字段：order\_id（订单 ID，可能重复）、user\_id（用户 ID）、order\_date（下单日期，格式为 “YYYY-MM-DD” 或 “MM/DD/YYYY”）、amount（订单金额，部分含 “¥” 符号）、payment\_method（支付方式：“支付宝”“微信”“银联”）、status（订单状态：“已完成”“取消”“退款”）。请完成以下操作：

1.数据读取与预处理：

读取数据，删除order\_id完全重复的行（仅保留 1 行）；

统一order\_date格式为 “YYYY-MM-DD”，并转换为 datetime 类型。

2.缺失值与格式处理：

处理user\_id缺失：用 “未知\_”+order\_id的格式填充；

清洗amount：去除 “¥” 符号并转换为数值型，缺失值用该列的中位数填充。

3.异常值处理：

删除amount≤0 的行；

标记 “金额＞10000 元” 的订单为 “高额订单”（新增is\_high列，值为 True/False）。

4.统计分析：

按payment\_method分组，计算各支付方式的订单总数、总金额；

按status分组，计算各状态订单的平均金额；

统计每个月的总订单数（提示：从order\_date提取月份）。