

14 数据ETL作业 – part-2

▶ 数据文件：

- ▶ pharmacy_data.csv

▶ 数据处理要求：

- ▶ (1) 从csv文件将数据读入数据框；
- ▶ (2) 用summary()函数查看数据集概况；
- ▶ (3) 重命名列名“日期”为“销售日”（可用rename()函数），并改数据类型为日期类型（可用ymd()函数）；
- ▶ (4) 删除日期和顾客编号数据缺失的数据行；
- ▶ (5) 数量、金额、折扣金额的缺失值分别用同品名的相应均值插补；
- ▶ (6) 数量为负数的行排除；
- ▶ (7) 所有数据按照销售日降序排列；
- ▶ (8) 增加折扣率新列；
- ▶ (9) 统计总实际销售额、总销售单数；
- ▶ (10) 统计各种商品的销售单数、总销售额、平均价格；
- ▶ (11) 按月统计各种商品的销售单数、销售额、平均价格
- ▶ (12) 统计每位顾客每月的实际总消费金额。