1. 缺失值
2. 异常值
3. 变量转化
4. 新变量生成
5. 数据清洗
6. 数据采样
7. 特征选择