【项目04】 视频网站数据清洗整理和结论研究

作业要求：

1、数据清洗 - 去除空值

要求：创建函数

提示：fillna方法填充缺失数据，注意inplace参数

2、数据清洗 - 时间标签转化

要求：

① 将时间字段改为时间标签

② 创建函数

提示：

需要将中文日期转化为非中文日期，例如 2016年5月24日 → 2016.5.24

3、问题1 分析出不同导演电影的好评率，并筛选出TOP20

要求：

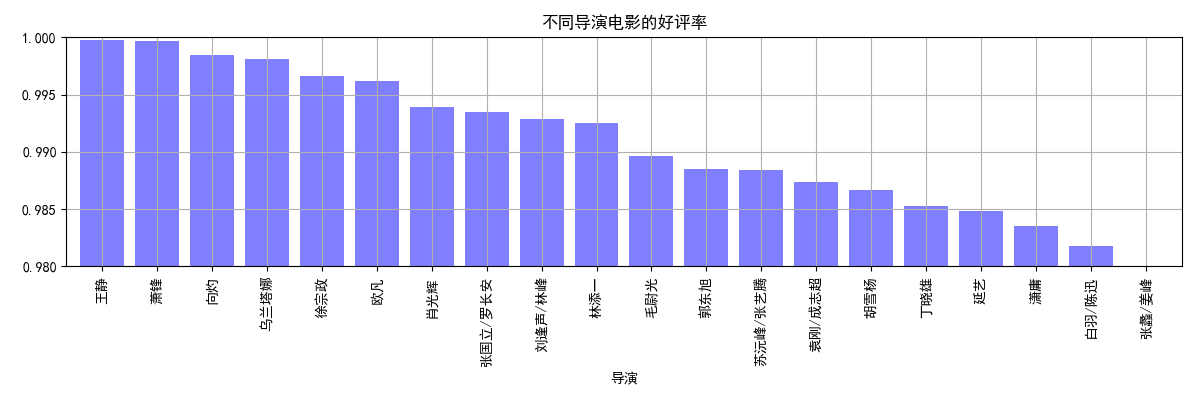
① 计算统计出不同导演的好评率，不要求创建函数

1. 通过多系列柱状图，做图表可视化

提示：

① 好评率 = 好评数 / 评分人数

② 可自己设定图表风格



4、问题2 统计分析2001-2016年每年评影人数总量

要求：

① 计算统计出2001-2016年每年评影人数总量，不要求创建函数

② 通过面积图，做图表可视化，分析每年人数总量变化规律

③ 验证是否有异常值（极度异常）

④ 创建函数分析出数据外限最大最小值）

⑤ 筛选查看异常值 → 是否异常值就是每年的热门电影？

提示：

① 通过箱型图验证异常值情况

② 通过quantile(q=0.5)方法，得到四分位数

③ IQR=Q3-Q1

1. 外限：最大值区间Q3+3IQR,最小值区间Q1-3IQR （IQR=Q3-Q1）
2. 可自己设定图表风格

