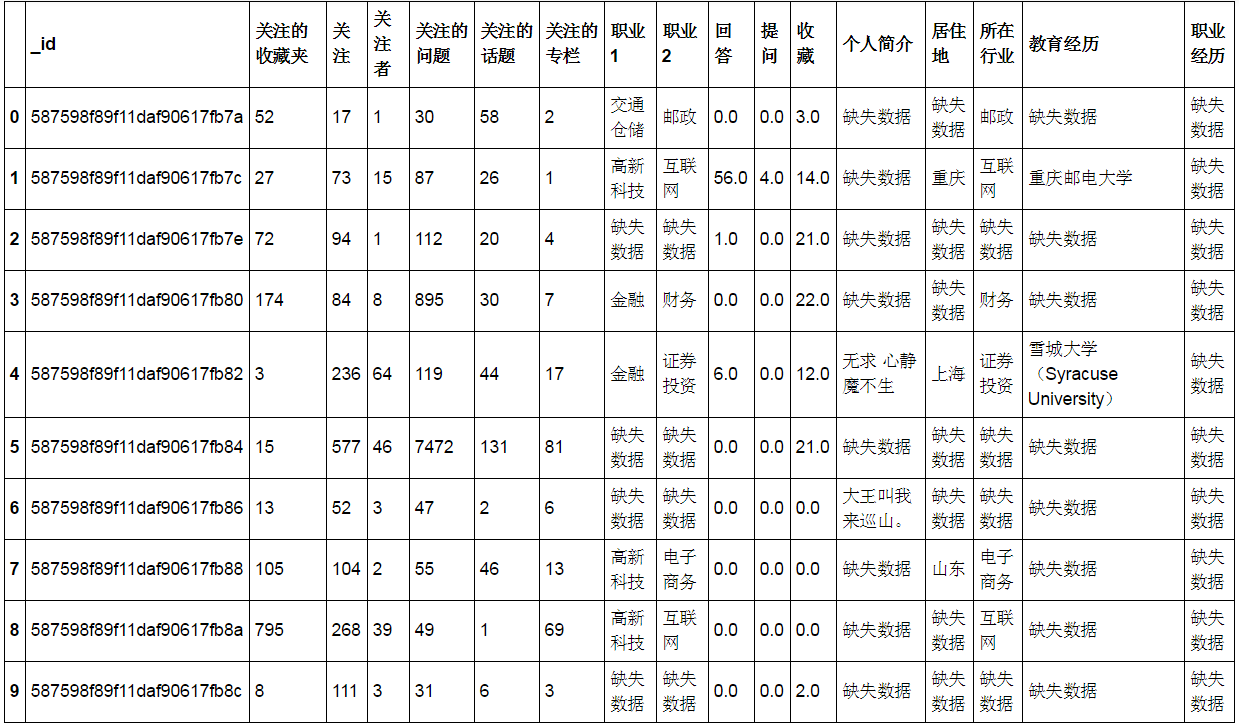
【项目03】 知乎数据清洗整理和结论研究

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、问题1 知友全国地域分布情况，分析出TOP20

要求：

① 按照地域统计 知友数量、知友密度（知友数量/城市常住人口），不要求创建函数

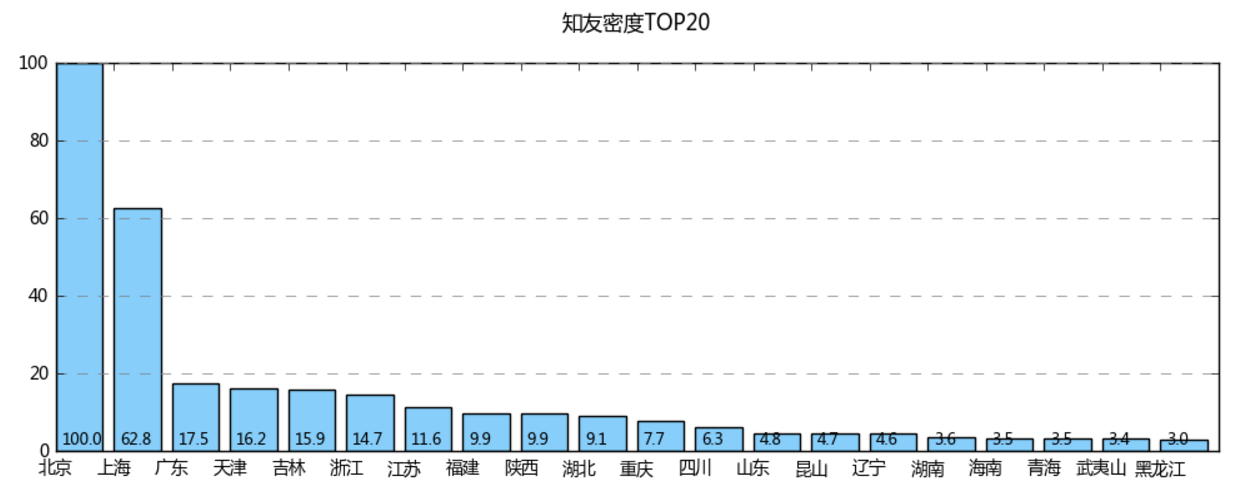
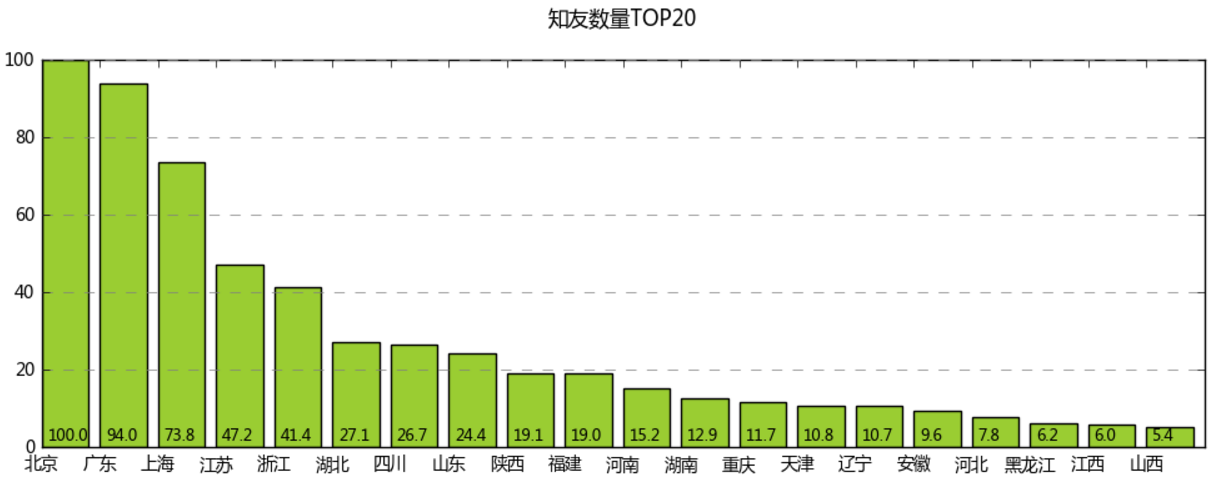
② 知友数量，知友密度，标准化处理，取值0-100，要求创建函数

1. 通过多系列柱状图，做图表可视化

提示：

① 标准化计算方法 = (X - Xmin) / (Xmax - Xmin)

② 可自行设置图表风格



3、问题2 知友全国地域分布情况，分析出TOP20

要求：

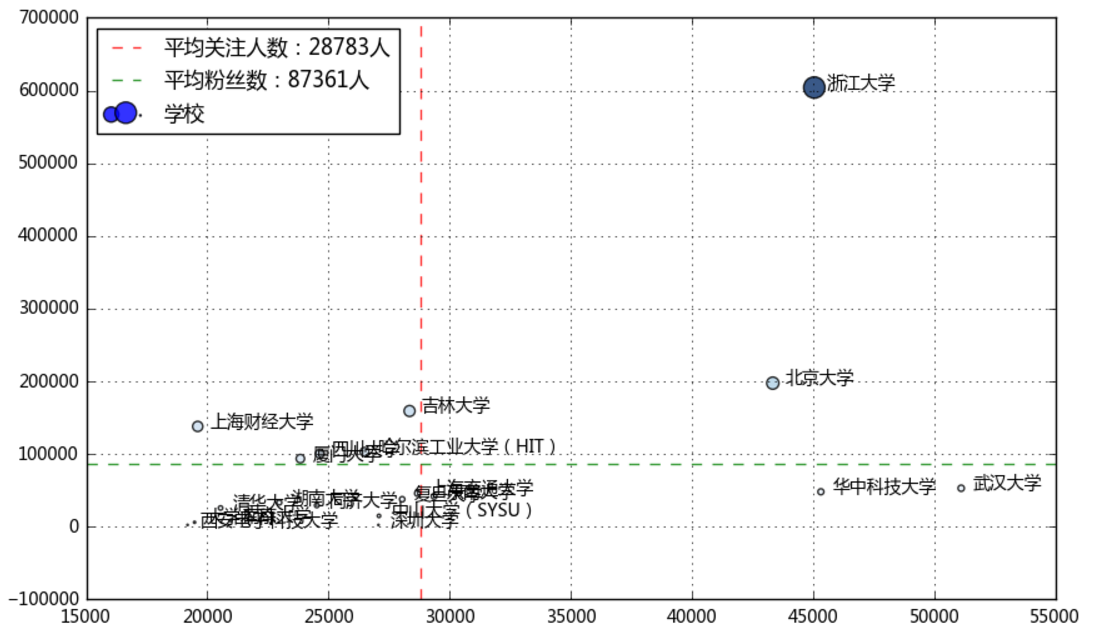
① 按照学校（教育经历字段） 统计粉丝数（‘关注者’）、关注人数（‘关注’），并筛选出粉丝数TOP20的学校，不要求创建函数

② 通过散点图 → 横坐标为关注人数，纵坐标为粉丝数，做图表可视化

③ 散点图中，标记出平均关注人数（x参考线），平均粉丝数（y参考线）

提示：

1. 可自行设置图表风格



广州：花城、羊城；重庆：山城；湘潭：莲城；长沙：星城、潭城；成都：蓉城、锦城；昆明：春城、花城；武汉：江城；金华：婺城；莆田：荔城；潮州：凤城；济宁：任城；徐州：彭城；大同：平城；嘉兴：禾城；安庆：宜城；西昌：月城；扬州：芜城；温州：鹿城；衢州：柯城；蚌埠：珠城；泉州：锦城；漳州：芗城；许昌：烟城；惠州：鹅城；柳州：龙城；泸州：酒城；内江：甜城；青岛：岛城；烟台：港城；曲阜：圣城；东营：油城；衡阳：雁城；福州：榕城；大连：滨城；长春：车城；苏州：水城；厦门：鹭城；鞍山：钢城；呼市：青城；哈尔滨：冰城；齐齐哈尔：鹤城；潍坊：风筝城；石狮：服装城；拉萨：日光城；绍兴：蠡城；济南：泉城、历城；南昌：洪城、英雄城；南京：石头城、金陵