

Lab4 Numpy & Matplotlib & Pandas 实践

实验简介：熟悉 NumPy、Matplotlib、Pandas 的基本操作，掌握其常用函数。

作业要求：仅提交代码文件（代码格式为“xxx.ipynb”，关键语句和逻辑要有注释），代码需要运行结果。

实验步骤：

- (1) 第 4 讲学习笔记
- (2) 操作说明以下 Numpy 函数的功能和用法：
 - 1) np.sort(); 2) np.argsort(); 3) np.tile(); 4) np.hstack(); 5) np.vstack();
 - 6) np.r_(); 7) np.c_(); 8) np.concatenate(); 9) np.newaxis(); 10) meshgrid()
- (3) 利用 Numpy 的分组及统计函数，计算四重奏数据(anscombe.csv)的统计值。
- (4) 利用 Pandas 的分组及统计函数，计算四重奏数据(anscombe.csv)的统计值。
- (5) 利用 Matplotlib 与 Seaborn，在给定数据集(shenzhenhouse.csv)上完成以下任务：
 - 1) 绘制直方图：房价单价分布；
 - 2) 绘制箱型图：各区房价分布、不同楼层房价分布；
 - 3) 绘制热力矩阵（用相应函数计算相关系数），查看不同变量之间的相关性强弱；
 - 4) 绘制累计频率直方图：房价单价分布；
 - 5) 通过可视化方法探索人们最关注什么户型的房子；
 - 6) 自定义不同的价格区间，绘制房子单价在定义的价格区间内分布的直方图。

参考链接：

Numpy: <http://www.numpy.org/>

Pandas Tutorials: https://pandas.pydata.org/pandas-docs/stable/getting_started/tutorials.html

Matplotlib <http://matplotlib.org/>

Matplotlib gallery <https://matplotlib.org/gallery.html>

Seaborn 0.9 中文文档 <https://www.cntofu.com/book/172/index.html>