

# 实验七指导书

---

## 实验要求

该实验为组队完成任务，组队人数要求**最多3人一组，可以单人一组，独立完成会有适当加分，实验截止到考试周**，时间较长，分数占比较大，希望好好对待。

实验要求每组成员对所给的数据或自己寻找的数据(**自己寻找的数据在提交时需一块提交**)进行分析，分析内容可以有文字描述及图表分析，为降低实验难度，**允许使用excel进行图表绘制**，但有能力的情况下尽可能采用代码绘制。

实验为开放性实验，没有统一答案标准，分析合理即可。

**Tips:** 任何一个程序语言都存在许多功能库，不建议对功能库的所有内容都进行学习（因为你根本都学不完，这也是不现实的），只需要对你用到的函数进行简单的学习即可，做到会用就行。提供的示例代码中有使用的例子，学习的第一步就是学习“抄”，学会模仿。另外，！！**学习使用百度**对你所遇到的问题进行检索，**这很重要**，互联网就是最好的老师！！

## 数据下载

**实验示例代码及数据下载地址：**<https://icloud.qd.sdu.edu.cn:7777/link/128486E768FA1BF68E40E1443B2789ED>

下载链接中每个数据文件夹中均包含任务数据、数据说明文档、示例程序代码。

部分文件因为编码问题可能无法直接打开查看，可用pandas读取阅读。

需要注意的示例代码中有些库和函数在课程中并未提及，需要自行安装学习（**助教对这些库也不了解**）。

另外给大家提供了去年优秀数据分析作业作为参考，**不可抄袭**。

## 报告提交

实验七不进行验收，最后**各组提交**一个压缩包即可，命名格式为：**组员成员姓名-实验七**

报告的提交时间截止到6月26日晚12:00，提交到山大网盘：<https://icloud.qd.sdu.edu.cn:7777/link/56F0F2E3A35C8C79A2AD260DD66EF940>

压缩包需要包含以下内容

- 实验报告。命名格式及文件形式为：**组员成员姓名-实验七.pdf**；实验报告内容需要有
  - 需要注明组员成员姓名及学号
  - 对数据的分析，需要有文字描述及可视化图片（**不允许截代码充当图片，也不允许粘贴大段代码充当文字描述，可以带一两行代码加几句中文作为文字描述**）
  - **采用jupyter进行开发的可以直接导出pdf或html作为实验报告**
  - 实验报告没有模板要求，自行对内容进行排版，排版的好看一点
- 程序代码。带注释，需要保证可复现。
- 所使用的数据（**允许自己找外部数据**）

以上命名及文件格式，若不按要求来，均扣分。

## 课堂展示

有意向课堂上做展示的，可以在6月8日之前，提前联系助教。展示内容不要求程序完整，做多少展示多少，主要在于讲出自己组对数据是如何分析的，以及后续打算如何操作。展示方式可以在ppt和程序展示中进行选择。**展示加分**。展示时间根据报名组数另行通知。

## 附录

MacOS用户需要注意：有些示例程序在使用matplotlib、seaborn之类的画图程序时，可能会遇到中文乱码，具体解决方式可以参照以下链接

<https://www.jianshu.com/p/c855bae924ce>

<https://mathsyouth.github.io/2019/06/12/macOS-matplotlib>