# 项目 0: 入门与基础

## 预测泰坦尼克号乘客幸存率

### 安装要求

这个项目要求使用 \*\*Python 2.7\*\* 以及安装下列python库

- [NumPy](http://www.numpy.org/)

- [Pandas](http://pandas.pydata.org)

- [matplotlib](http://matplotlib.org/)

- [scikit-learn](http://scikit-learn.org/stable/)

​

你还需要安装和运行 [Jupyter Notebook](http://jupyter.readthedocs.io/en/latest/install.html#optional-for-experienced-python-developers-installing-jupyter-with-pip)。

优达学城推荐学生安装 [Anaconda](https://www.continuum.io/downloads)，一个包含了项目需要的所有库和软件的 Python 发行版本。[这里](https://classroom.udacity.com/nanodegrees/nd002/parts/0021345403/modules/317671873575460/lessons/5430778793/concepts/54140889150923)介绍了如何安装Anaconda。

如果你使用macOS系统并且对命令行比较熟悉，可以安装[homebrew](http://brew.sh/)，以及brew版python

```bash

$ brew install python

```

再用下列命令安装所需要的python库

```bash

$ pip install numpy pandas matplotlib scikit-learn scipy jupyter

```

### 代码

​

核心代码在 `titanic\_survival\_exploration.ipynb` 文件中，辅助代码在 `titanic\_visualizations.py` 文件中。尽管已经提供了一些代码帮助你上手，你还是需要补充些代码使得项目要求的功能能够成功实现。

### 运行

​

在命令行中，确保当前目录为 `titanic\_survival\_exploration/` 文件夹的最顶层（目录包含本 README 文件），运行下列命令：

```bash

$ jupyter notebook titanic\_survival\_exploration.ipynb

```

​

这会启动 Jupyter Notebook 把项目文件打开在你的浏览器中。

对jupyter不熟悉的同学可以看一下这两个链接：

- [Jupyter使用视频教程](http://cn-static.udacity.com/mlnd/how\_to\_use\_jupyter.mp4)

- [为什么使用jupyter？](https://www.zhihu.com/question/37490497)

​

​

​

​

​

​

​

​

​

​

​

​

​

​

### 数据

​

这个项目的数据包含在 `titanic\_data.csv` 文件中。文件包含下列特征：

​

- \*\*Survived\*\*：是否存活（0代表否，1代表是）

- \*\*Pclass\*\*：社会阶级（1代表上层阶级，2代表中层阶级，3代表底层阶级）

- \*\*Name\*\*：船上乘客的名字

- \*\*Sex\*\*：船上乘客的性别

- \*\*Age\*\*：船上乘客的年龄（可能存在 `NaN`）

- \*\*SibSp\*\*：乘客在船上的兄弟姐妹和配偶的数量

- \*\*Parch\*\*：乘客在船上的父母以及小孩的数量

- \*\*Ticket\*\*：乘客船票的编号

- \*\*Fare\*\*：乘客为船票支付的费用

- \*\*Cabin\*\*：乘客所在船舱的编号（可能存在 `NaN`）

- \*\*Embarked\*\*：乘客上船的港口（C 代表从 Cherbourg 登船，Q 代表从 Queenstown 登船，S 代表从 Southampton 登船）