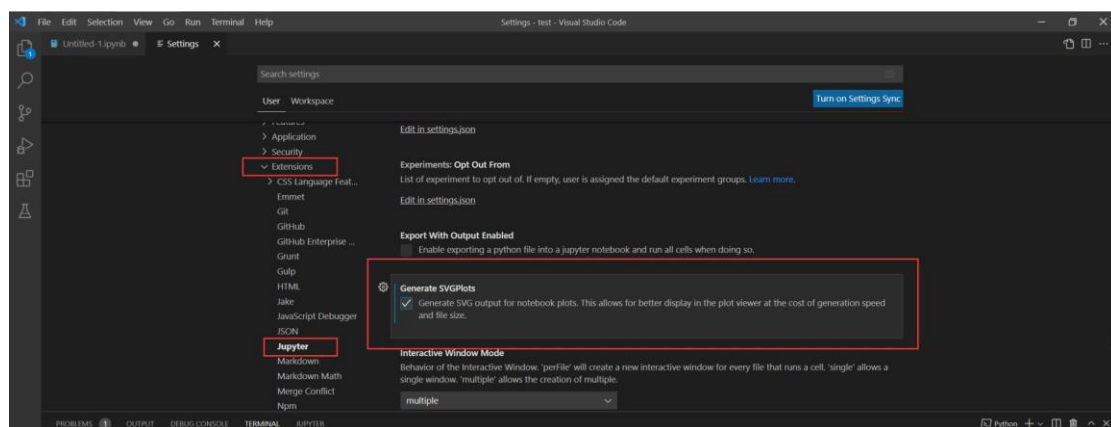


## 任务 3: Numpy、Pandas 与 Matplotlib

参考: <https://www.bilibili.com/video/BV1ZM4y1u7uF?p=5>

- Numpy array 以及围绕 array 的一些基本操作方法, 比如 shape、reshape、where、通过[:]的切片
- Pandas 中 Dataframe 和 Series 的基础概念、赋值与枚举
- Pandas 的文件读取, 尤其是 CSV
- Pandas 缺失值处理
- Pandas 中的 DataFrame Merge 与 Grouping
- Pandas 与 Matplotlib.pyplot 的联动, 使用 DataFrame 数据绘图
- Matplotlib.pyplot 需要学员掌握 plot, bar, scatter, figure & subplot, text, labels, axis & grid, legend 等的用法
- Matplotlib 需要额外掌握 image(imshow), pie

附: 提高生成图像清晰度的小技巧



在 VSCode 的设置 - Extensions - Jupyter 中，勾选 Generate SVGPlots。

## 完成任务：

### 一、 根据 `chipotle.tcv` 编程来完成以下任务：

1. 导入数据
2. 查看前 10 行内容
3. 打印出该数据的列数
4. 打印出全部列的名称
5. 找出下单数最多的 item
6. 打印出 item\_name 的数量
7. 找出在 choice\_description 中，下单次数最多的商品
8. 打印出商品被下单的总数
9. 打印出在该数据集对应的时期内的收入(revenue)数额
10. 打印出在该数据集对应的时期内的订单(order)数
11. 打印出每一笔订单的平均总价

### 二、 根据 `titanic.csv` 完成如下任务：

1. 导入数据
2. 将 PassengerId 设置为索引
3. 绘制一个展示男女乘客比例的扇形图
4. 绘制一个展示船票 Fare, 与乘客年龄和性别的散点图
5. 绘制一个展示船票价格的直方图
6. 有多少人生还？

### 三、 根据 `Apple_stock.csv` 完成如下任务：

- 1、导入数据
- 2、读取数据并存为一个名叫 `apple` 的数据框
- 3、查看每一列的数据类型
- 4、将 `Date` 这个列转换为 `datetime` 类型
- 5、将 `Date` 设置为索引
- 6、查看是否有重复的日期吗？
- 7、将 `index` 设置为升序
- 8、找到每个月的最后一个交易日(business day)
- 9、数据集中最早的日期和最晚的日期相差多少天？
- 10、按照时间顺序可视化 `Adj Close` 值