# Markdown 文件批量转换为 pdf



用十几行 Python 代码和格式转换界黑魔法 Pandoc ，迅速搞定。

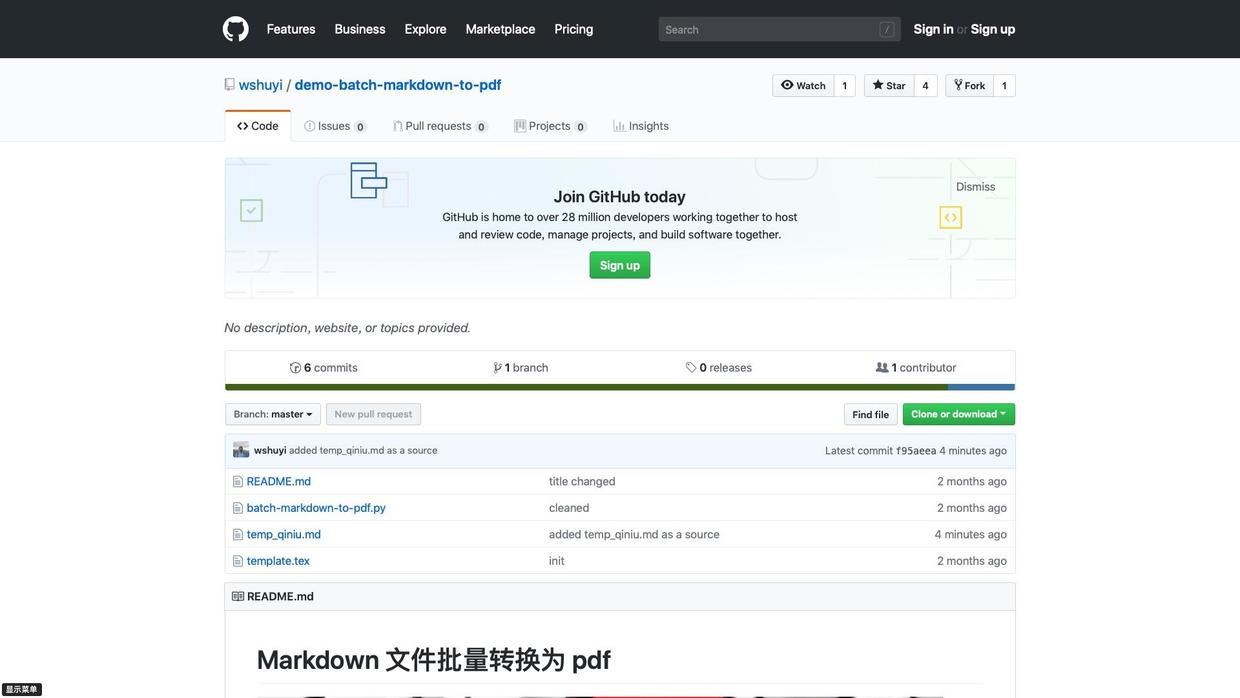
# 需求

有个朋友提出，希望把目录中的许多 markdown 文件，批量转换为对应名称的 pdf 格式文件。我于是编写了一个 Python 脚本，并且分享给你。如果你有类似的需求，欢迎使用。

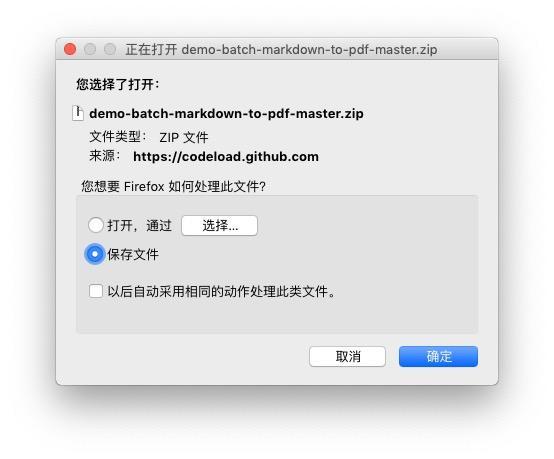
由于使用了 pandoc 作为转换工具，因此 Markdown 文件里的图片链接，不论是本地存储的（只测试了绝对路径情况），还是图床上的，都可以正确转换并且显示到 pdf 文件里。

# 数据

我已经把代码和样例 Markdown 文件，都为你放在了[这个 github repo](https://github.com/wshuyi/demo-batch-markdown-to-pdf)中。



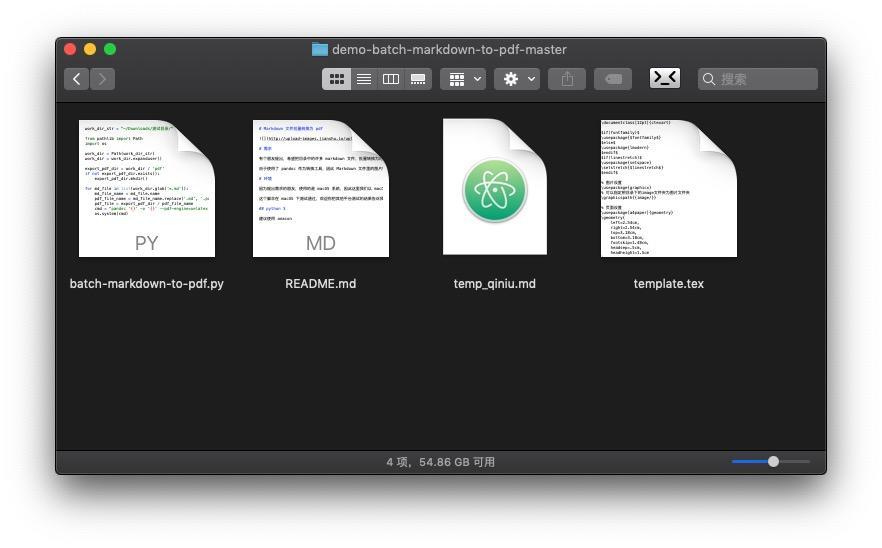
你可以直接点击[这个链接](https://github.com/wshuyi/demo-batch-markdown-to-pdf/archive/master.zip)，下载压缩包 demo-batch-markdown-to-pdf-master.zip。



在 macOS 上默认的下载位置，是 ~/Downloads。

下载后，解压该压缩包，咱们的演示目录就准备好了。名称是 ~/Downloads/demo-batch-markdown-to-pdf-master 。

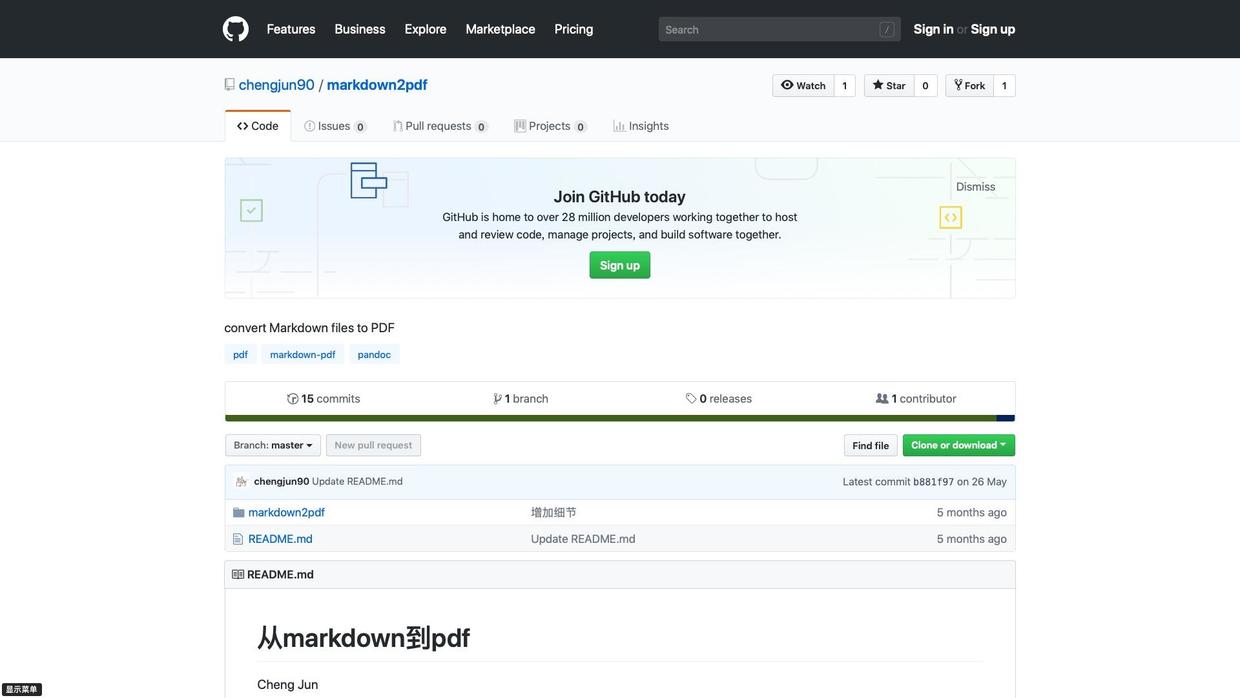
压缩包里面，有4个文件。



其中的batch-markdown-to-pdf.py是运行脚本；

temp\_qiniu.md 和 README.md 是咱们的两个示例 Markdown 文件。你尝试之后，可以换成自己的一批 Markdown 文件。

template.tex是转换是采用的模板，这个模板并非我做的，它来自于[这个](https://github.com/chengjun90/markdown2pdf) github 项目。



如果你对 latex 有研究，可以自行修改 template.tex 的内容，以控制输出 pdf 的样式。

# 环境

因为提出需求的朋友，使用的是 macOS 系统，因此这里我们以 macOS 系统的安装方式为准。注意下述工具实际上都是**跨平台**的。因此如果你使用的是 Windows 或者 Linux ，理论上也都是可以使用的。

这个脚本在 macOS 下测试通过，欢迎你把其他平台测试的结果告诉我。

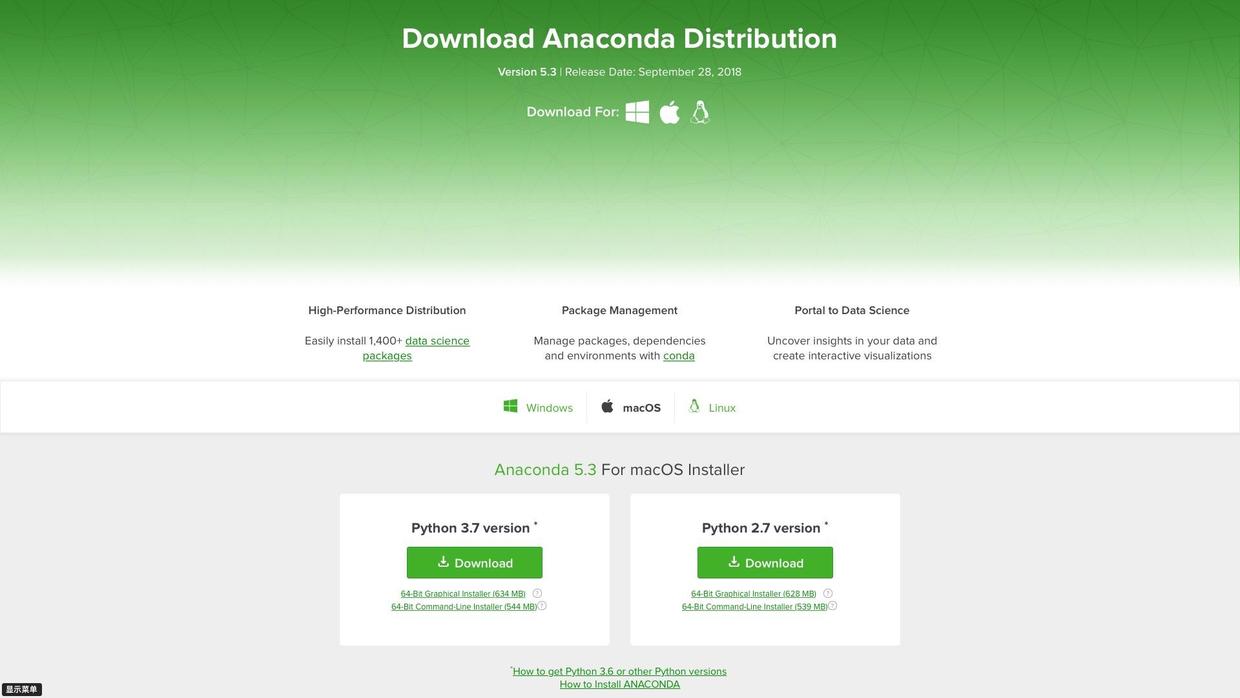
## python 3

在 macOS 上面安装 Python 3 ，有两种方式。

一种是安装 Anaconda 套件，另一种是使用 Homebrew 。

我们先说 Anaconda 套件安装方式。推荐普通用户使用。它不仅包含 Python 本身，还提前为你安装好了许多常用的依赖套件。

请到 [这个网址](https://www.anaconda.com/download/) 下载Anaconda的最新版本。

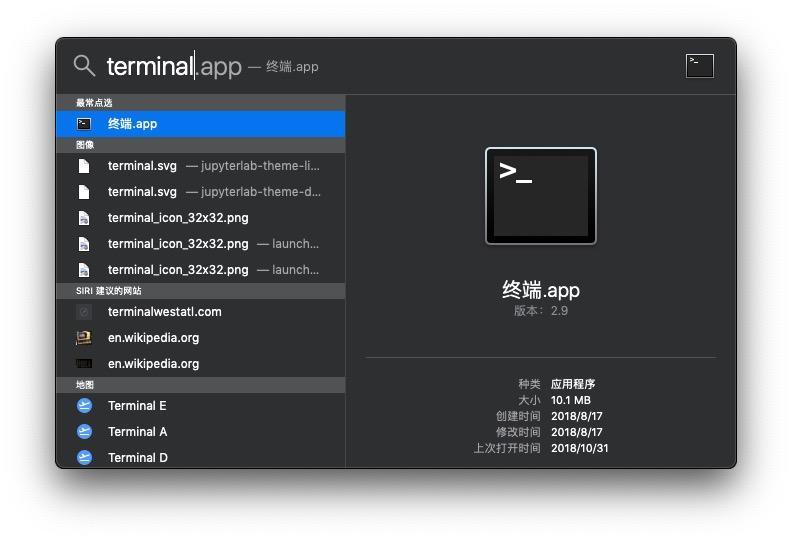


网站会主动识别你目前使用的操作系统。确定无误后，请选择左侧的 Python **3.7** 版本下载安装。

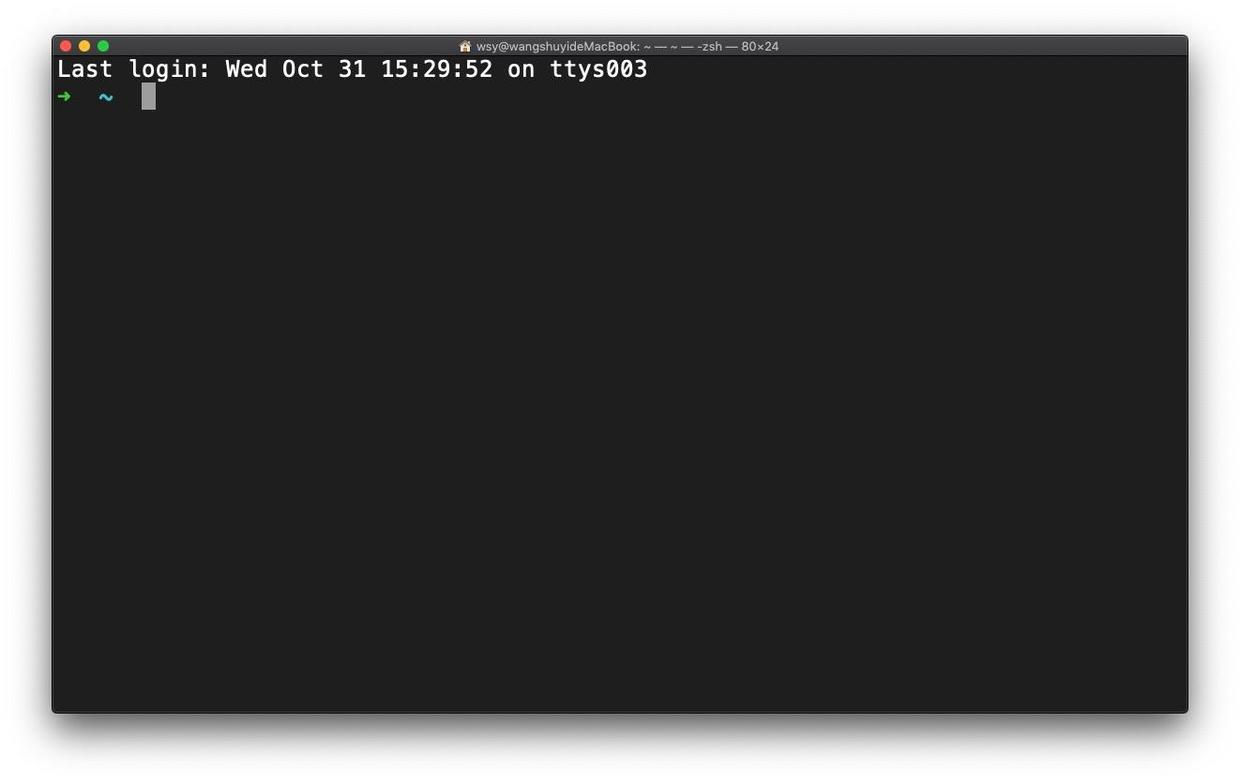
在 macOS 环境中，你下载下来的，是一个以 pkg 为扩展名的软件安装包。双击它，根据提示一步步前进就可以了。

安装完毕后，请打开一个终端窗口。

方法是在“聚焦搜索”(Spotlight)中，输入 Terminal.app 。



然后，回车就可以了。

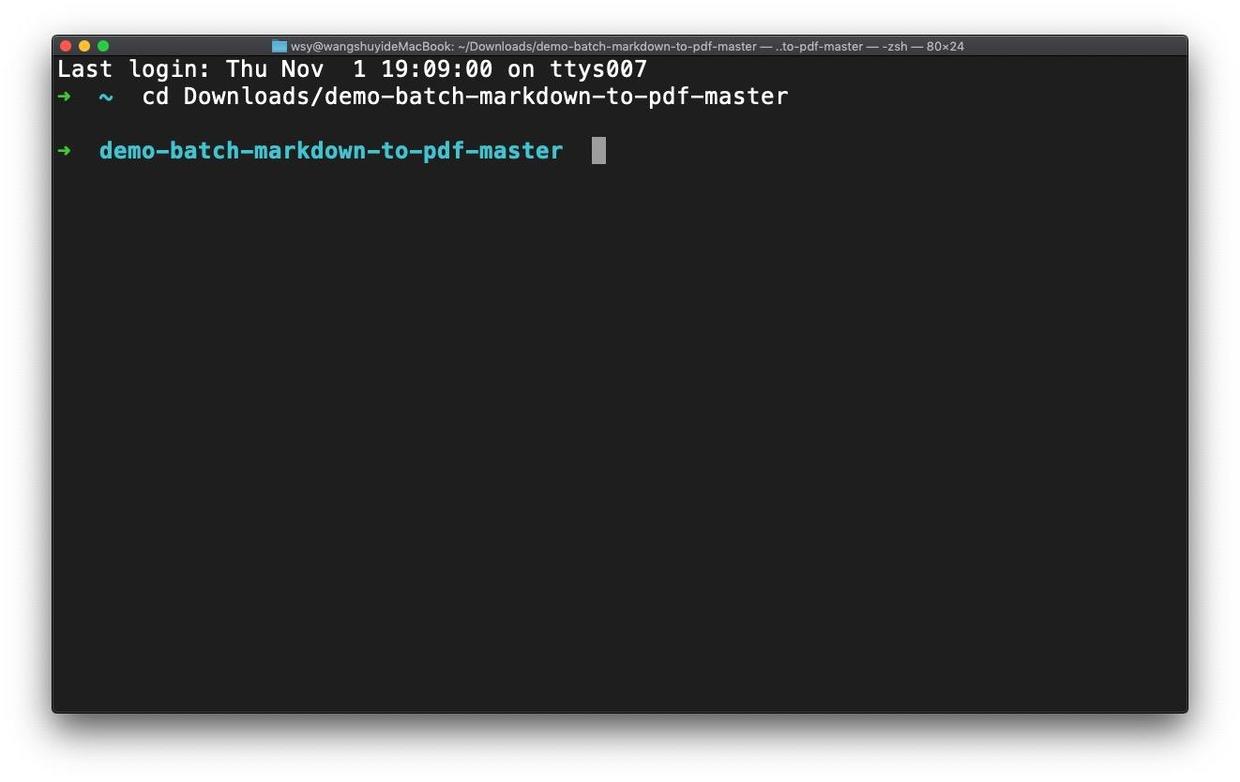


此时你会看到一个 ~ 提示符，这说明终端默认的初始位置，是用户的**家目录**。

咱们的演示目录位置位于 ~/Downloads/demo-batch-markdown-to-pdf-master ，所以你可以使用：

cd Downloads/demo-batch-markdown-to-pdf-master

这个命令，进入咱们的演示目录。



当你看到前面的路径提示，已经变成了 demo-batch-markdown-to-pdf-master ，就说明你已经定位到演示目录了。

对于高级用户，如果你觉得 Anaconda 安装了许多你不需要用到的软件包，那么也可以尝试 Homebrew 的安装方法。

首先你需要安装 XCode。安装方法请参见[这个链接](https://aaaaaashu.gitbooks.io/mac-dev-setup/content/XCode/index.html)。

然后，在终端窗口里面输入：

ruby -e "$(curl -fsSL https://raw.githubusercontent.com/Homebrew/install/master/install)"

之后，把下面这一条语句，添加到你的 ~/.profile 文件末尾：

export PATH=/usr/local/bin:/usr/local/sbin:$PATH

保存退出，新开一个窗口。

此时 Homebrew 已经安装好了，你可以执行以下命令安装 Python 3：

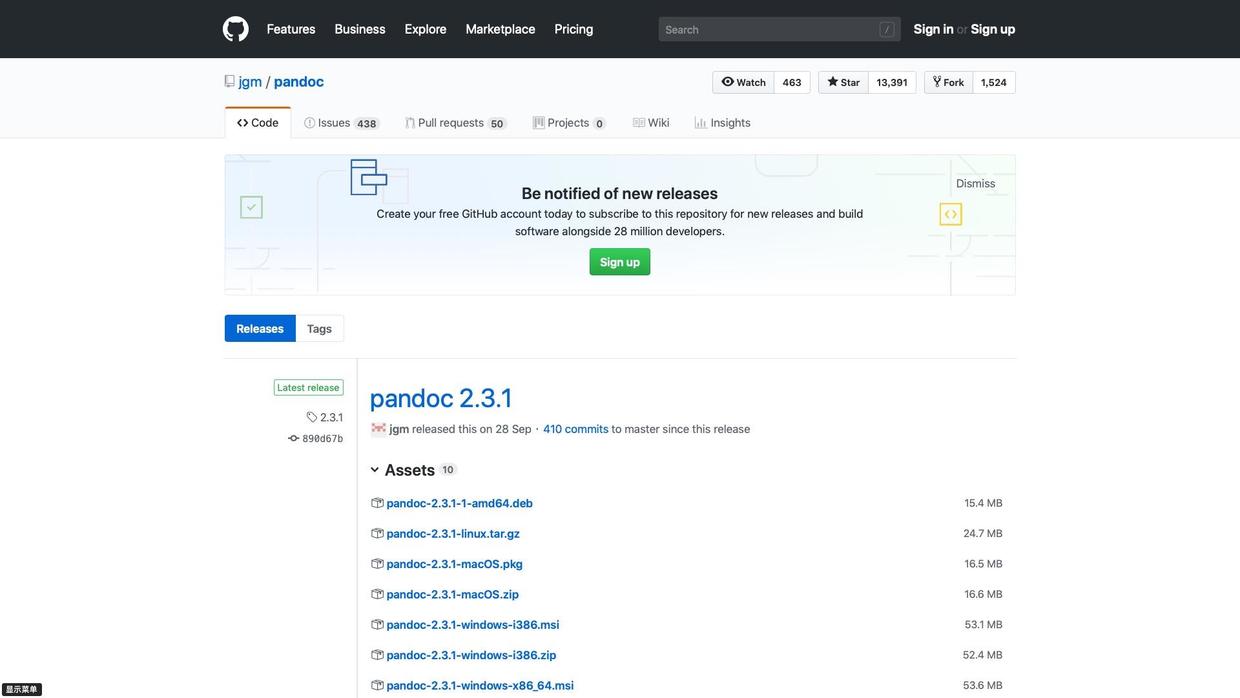
brew install python

之后，同样在终端中使用以下命令进入演示目录：

cd Downloads/demo-batch-markdown-to-pdf-master

## pandoc

请到[这个链接](https://github.com/jgm/pandoc/releases)，下载符合你使用操作系统的最新版本 pandoc ，并且进行安装。



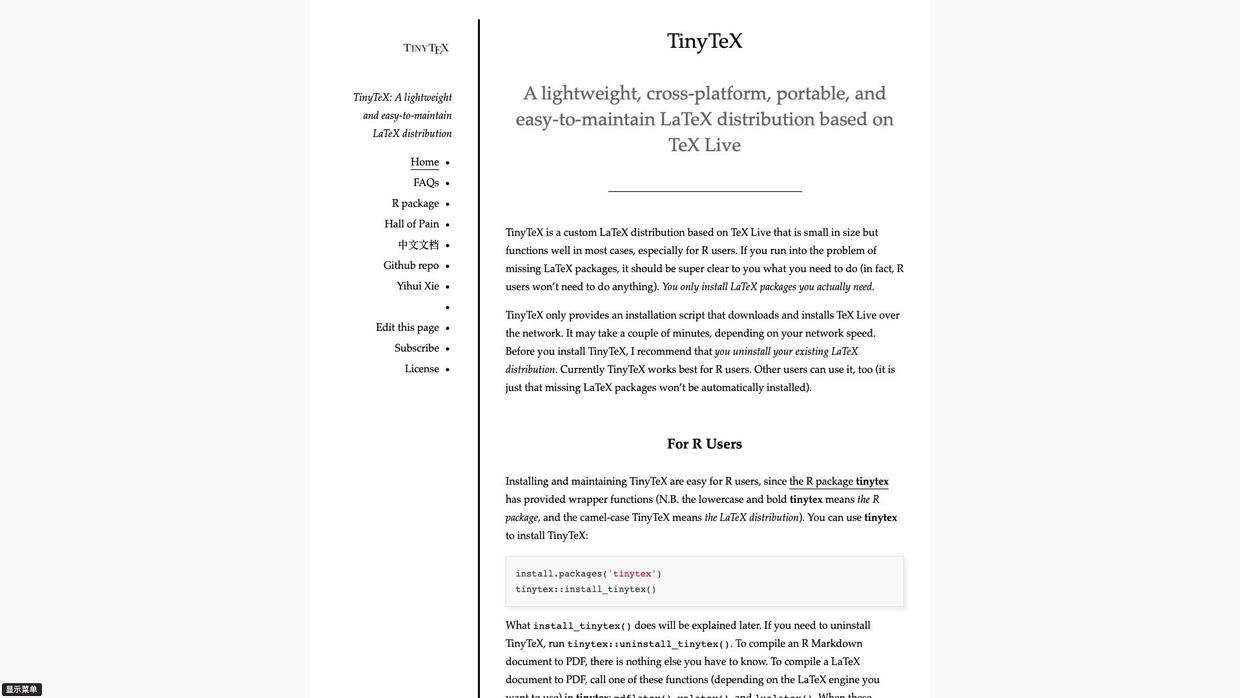
根据我们的情况，选择的就是 pandoc-2.3.1-macOS.pkg 。

下载下来的，依然是 pkg 安装包，还是双击，就可以根据提示安装了。

## tinytex

因为需要转换的 markdown 文件，大部分都是中文文档，因此转换到 pdf 的时候，需要 xelatex 的支持。

xelatex 可以用各种 latex 集成包来安装使用，例如 texlive 等。但是这里推荐谢益辉的 [tinytex](https://yihui.name/tinytex/) 包，简单小巧。



不过使用之前，建议删除掉系统里面原有的 texlive 等包。否则可能会造成冲突。

在终端窗口下，执行这个命令：

curl -sL "https://yihui.name/gh/tinytex/tools/install-unx.sh"

tinytex 就安装好了。

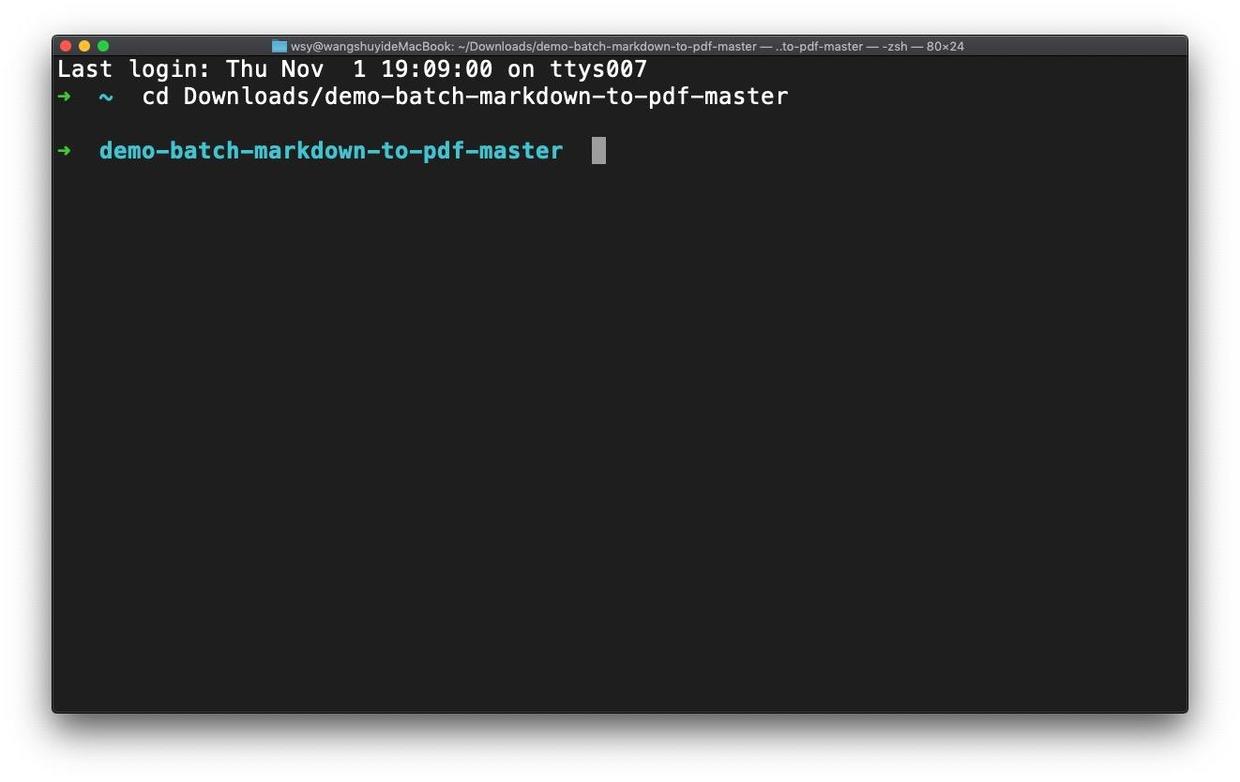
之后，为了能够更好地辅助我们进行转换，需要执行下列命令，安装扩展：

tlmgr install unicode-math filehook xecjk xltxtra realscripts fancyhdr lastpage ctex ms cjk ulem environ trimspaces zhnumber collection-fontsrecommended

好了，至此准备工作结束，我们该开始执行命令了。

# 运行

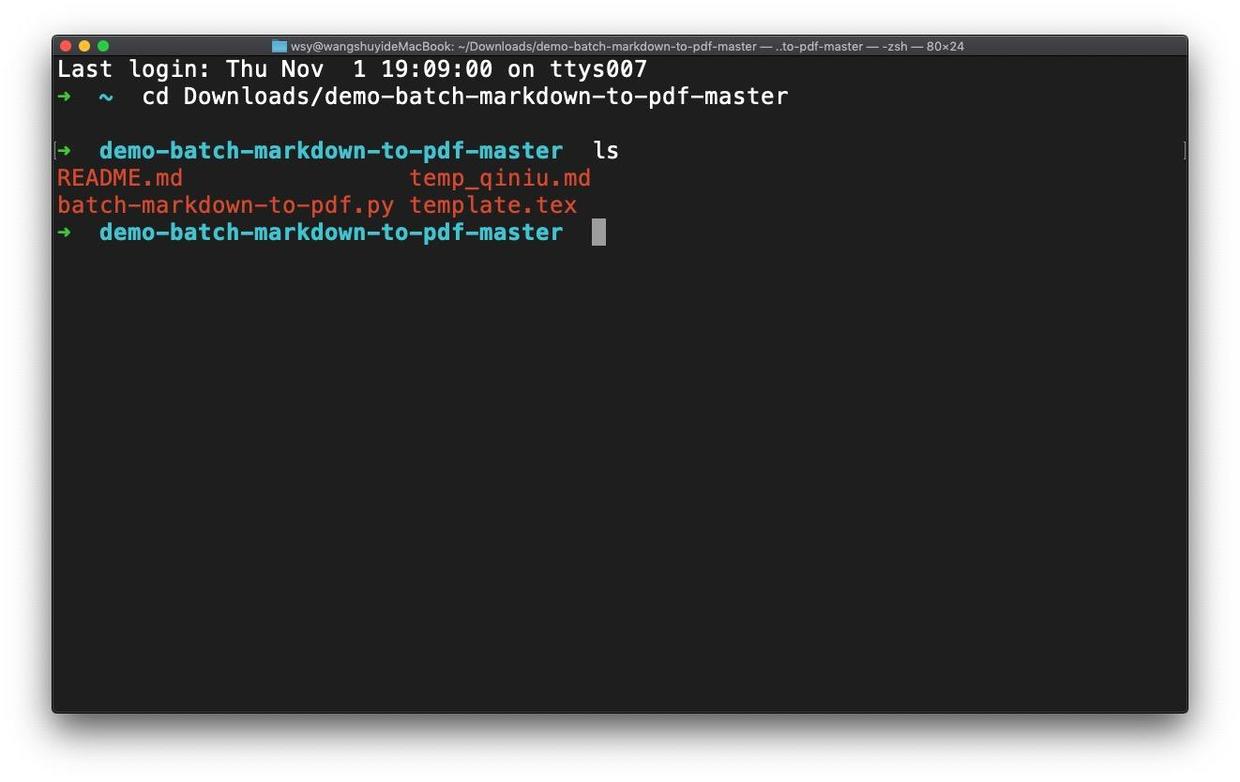
再次确认，你的终端下所在位置，为 demo-batch-markdown-to-pdf-master 。



执行目录查看命令：

ls

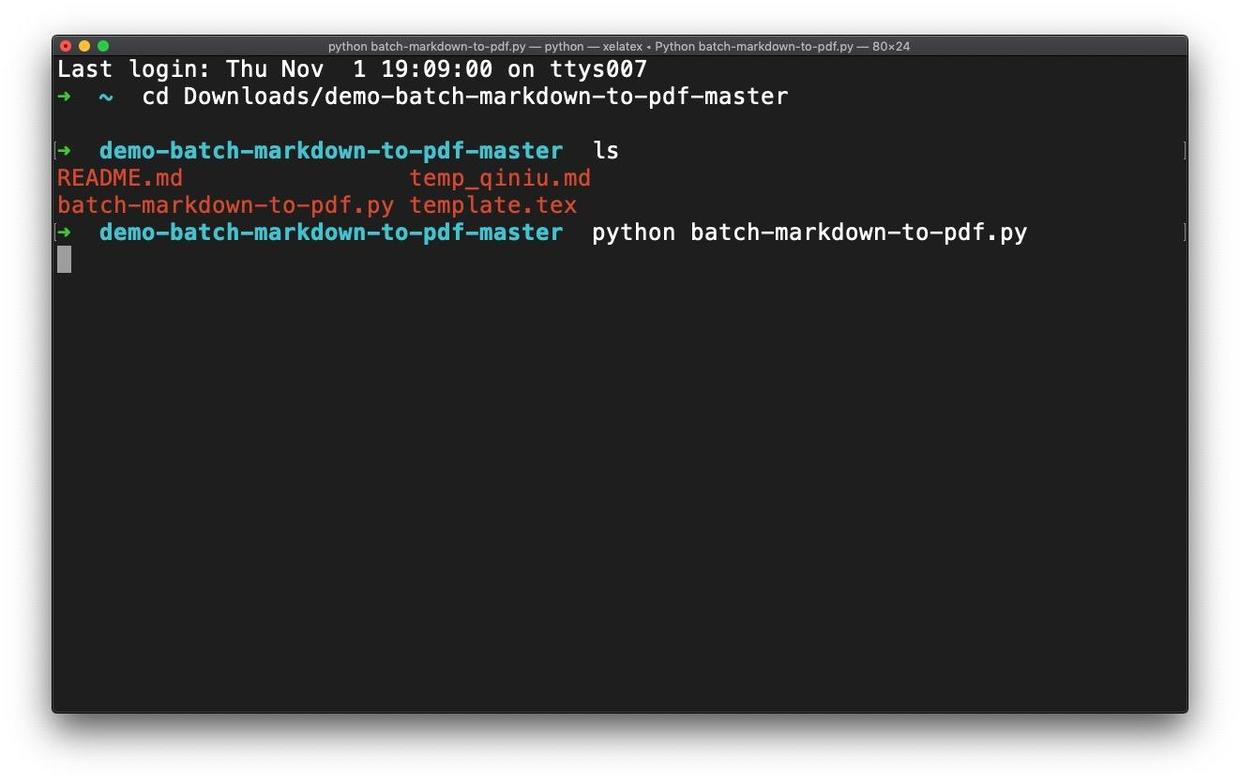
如果你看到返回的是如下信息，证明一切正常。



下面执行：

python batch-markdown-to-pdf.py

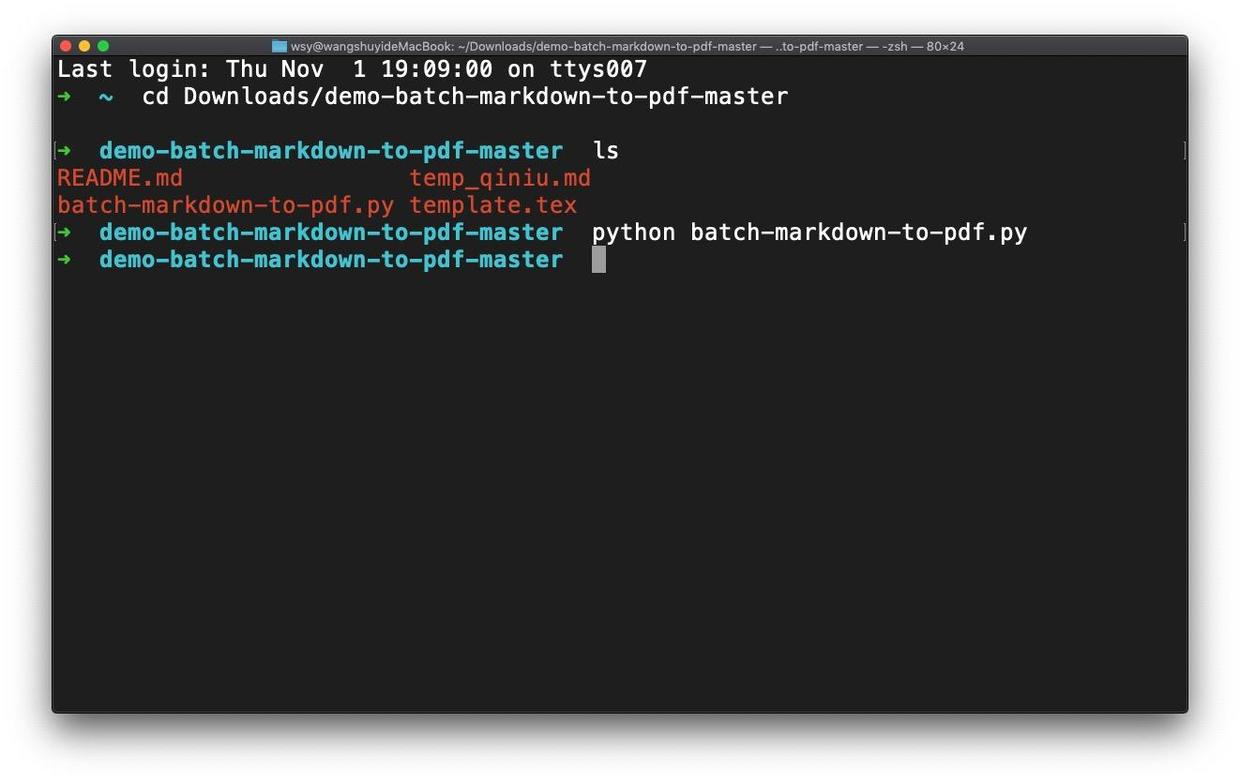
如果顺利，你会看到程序在运行，不过没有什么输出提示的。



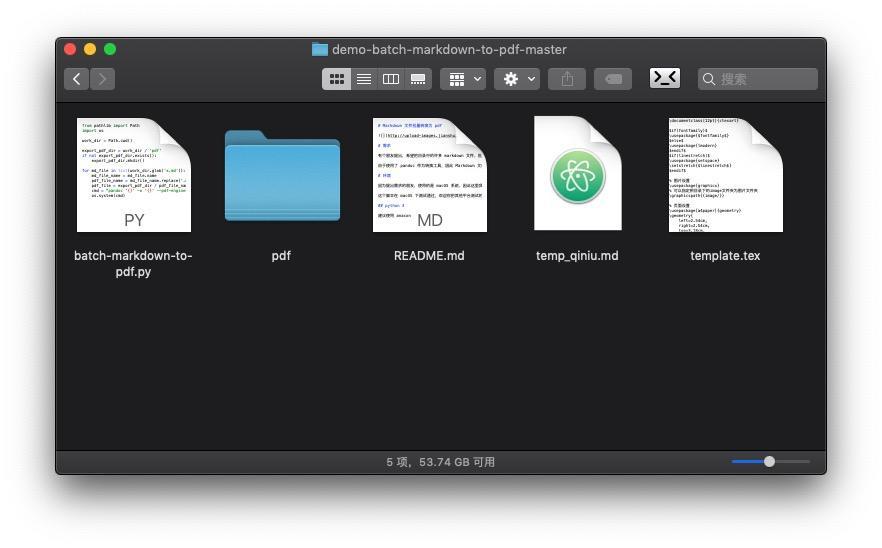
因为转换 pdf 的工作需要一些时间。所以如果你的 Markdown 文件很多，可能需要等一会儿。

请不要着急。去喝杯茶，看看书，休息一下。

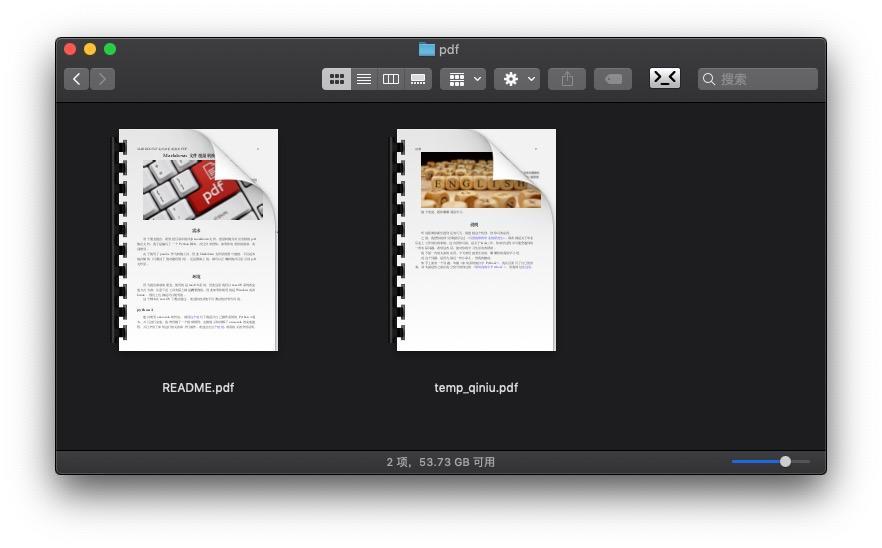
当你回来的时候，（但愿）已经转换完毕了。



回到“访达”（Finder） ，在我们的演示目录（~/Downloads/demo-batch-markdown-to-pdf-master）下面，你会看到新生成了一个文件夹，叫做 pdf 。



你的转换后 pdf 文件，应该已经在里面了。



双击打开，看看效果：



如果遇到问题，欢迎反馈给我。

祝使用愉快！

喜欢请点赞和打赏。还可以微信关注和置顶我的公众号[“玉树芝兰”(nkwangshuyi)](http://oejqwrqkh.bkt.clouddn.com/2016-10-11-22-26-16.jpg)。

如果你对 Python 与数据科学感兴趣，不妨阅读我的系列教程索引贴《[如何高效入门数据科学？](https://www.jianshu.com/p/85f4624485b9)》，里面还有更多的有趣问题及解法。