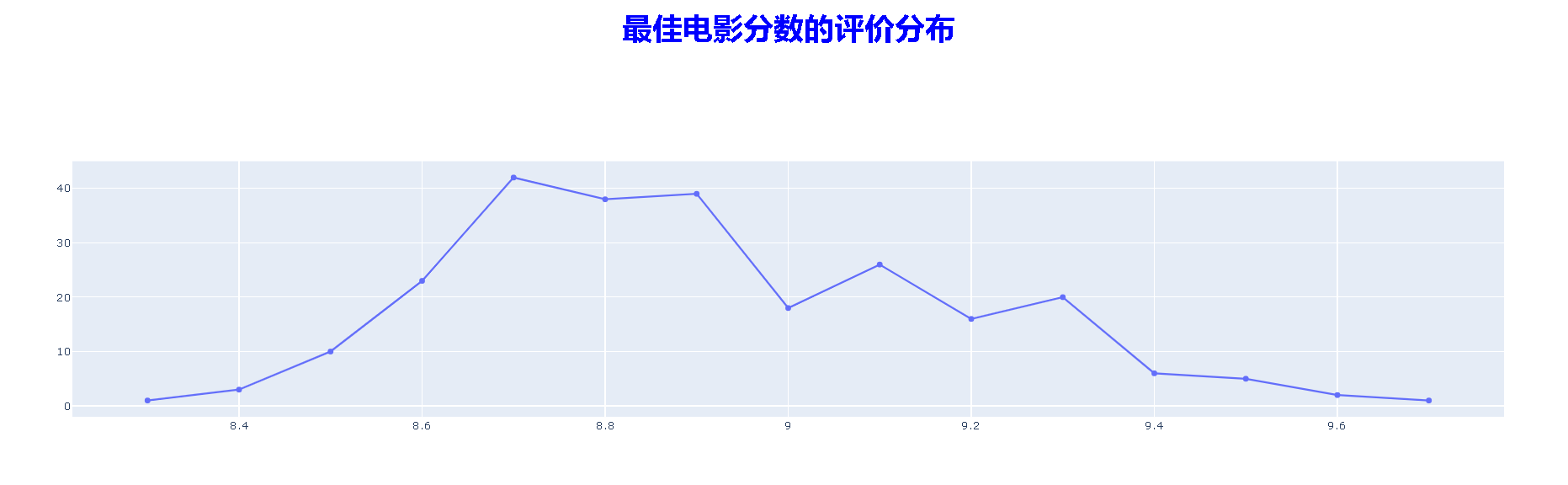
用visual studio code打开当前文件夹下的main.ipynb，假如有模块缺失，使用pip install下载相应的库（注意使用了Jupyter），运行所有单元，在运行到最后的app.run\_server()后，打开<http://127.0.0.1:8050/> 查看实验效果即可

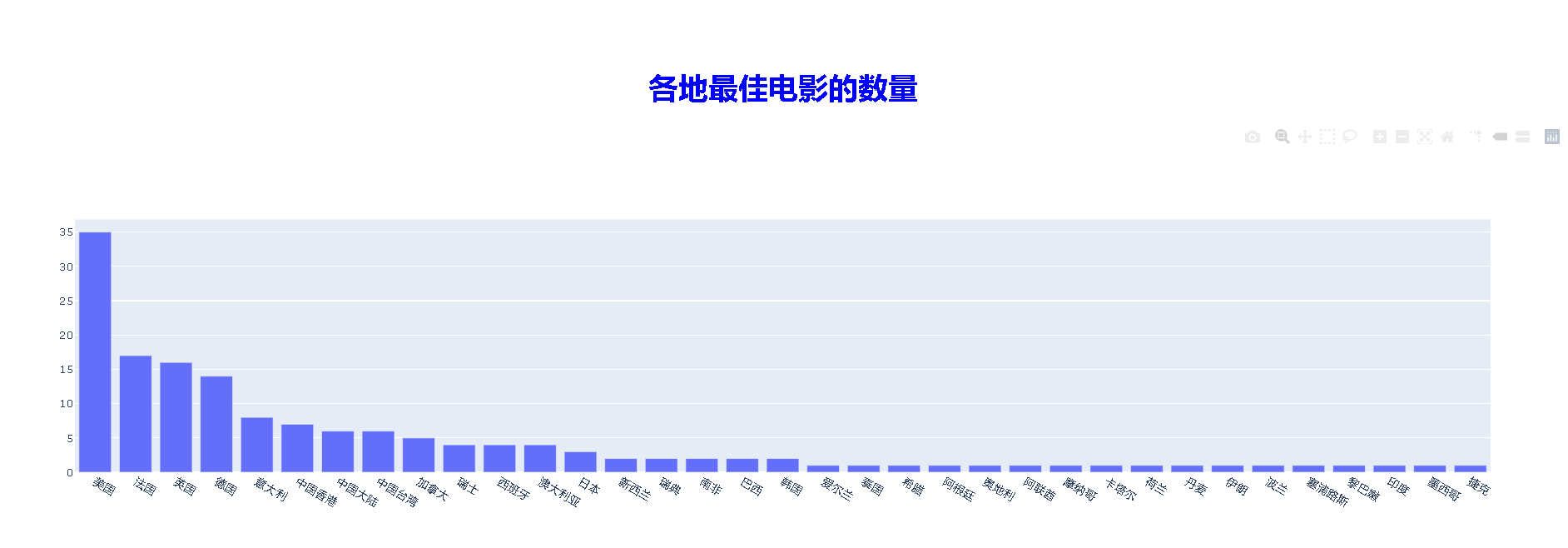
另外，虽然目标数据已经爬取完成，但也可以再次进行爬取，那就执行store.py，自动用selenium打开google浏览器，爬取到的图片部分（电影海报）会保存到sources的pics文件夹里，同时所有图片压缩存至“图片总览.pdf” ，另外，我根据这每隔4年的电影及其评分做了一个动态图，就是同级目录下的movies.gif

首先是豆瓣电影分数的评价分布



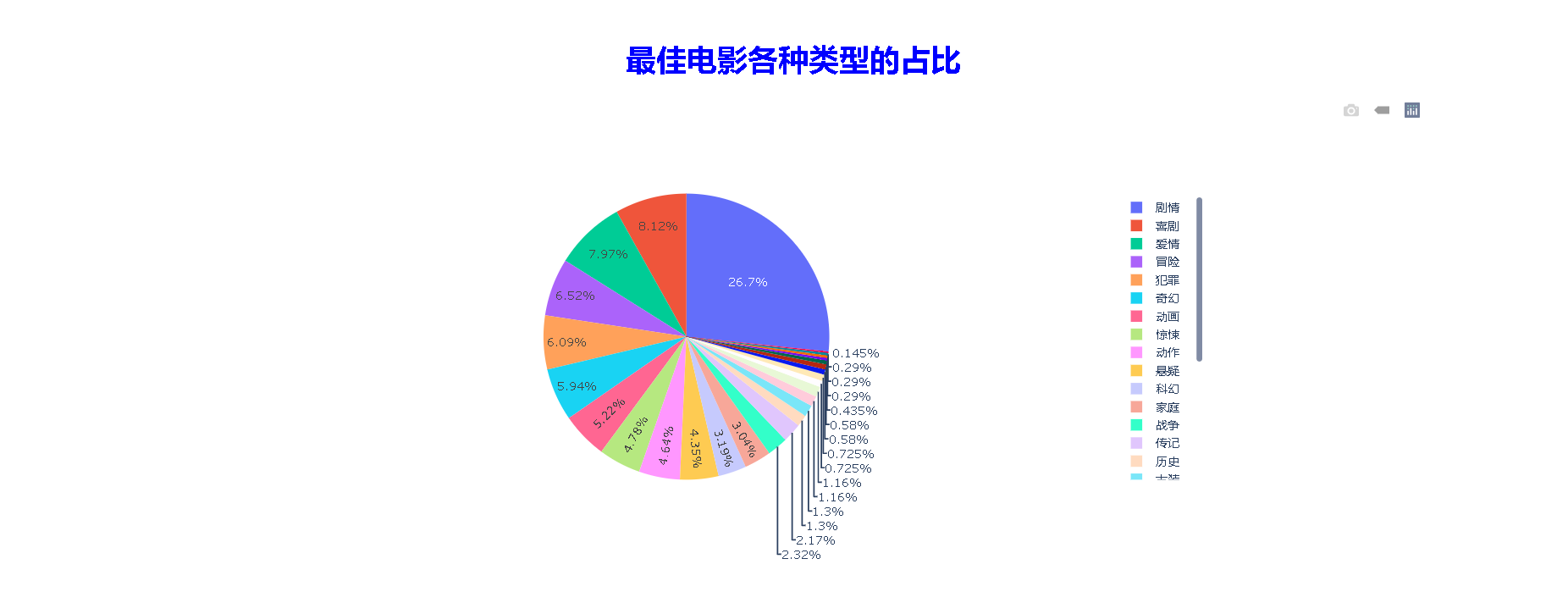
由此图可以得到豆瓣历年电影评分的分布，可以看到最佳电影评分集中在8.7到8.9之间

接下来是各地最佳电影的数量



最多的是美国，法国和英国，中国也在前列（而且还分地区）

这是最佳电影各种类型的占比



可以看到最多的是剧情片，其次是喜剧和爱情，剧情占据主导地位。

这是每4年出现的最佳电影及其评价，这里采用了pandas的降采样，可视化时采用了dash的slider



假如要查看所有电影的海报的话，可以打开同级目录下的“”图片总览.pdf“”