一：py文件功能介绍：

1. Scrapy58921.py：

从58921.com电影网站上爬取2015年至2018年数据并按照规定格式存储到MySQL指定数据库中

1. labvisual.py ：

在数据爬取并存储完毕后，可以对存储在数据库中的电影数据进行排序以及可视化处理，pyecharts可视化自动生成的html分别为：

1516\_month\_boxoffice.html

1718\_month\_boxoffice.html

根据用户输入内容，自动生成的html分别为：

（用户输入）年top（用户输入）词云.html

top（用户输入）电影题材.html

top（用户输入）劳模演员.html

3.movies\_gui文件： 使用tkinter工具生成了一个可以账号密码登陆的图形界面

二：数据库注意事项

1.连接数据库具体方法：

在各py文件中

将代码db = pymysql.connect(host='localhost', port=3306, user='\*\*\*\*\*', passwd='\*\*\*\*\*', db='movies\_database', charset="utf8")

\*\*\*\*\*处添加自己MySQL的账号以及密码

并手动建立名为movies\_database的数据库

或者在-仓库中利用已有数据库movies\_database\_movies\_table.sql

（密码为Oracle19990112）

2.数据库中movies\_table表形式按照如下方式构造



三：报表相关工具

用户需要自行下载安装phantomjs以及wkhtmltopdf工具

并且将bin文件的路径添加到计算机的环境变量中

phantomjs用于将html文件转为png文件

wkhtmltopdf用于将html文件转为pdf文件