数据库数据获取的架构思路

2018年9月16日

16:12

**1.门票信息爬取**

1.首先连接数据库,获取所有城市信息,然后使用这条url爬取到全中国所有城市的景点信息,爬取景点下面的门票内容存入数据库中,门票所属景点名设为对应的景点名

url:

<http://piao.ctrip.com/dest/u-_b3_a4_c9_b3/s-tickets/?searchValue=%E9%95%BF%E6%B2%99>

计算机生成了可选文字:
名称
特惠票
特惠票
名称
特惠票
特惠票
点评
长沙生态动物园
〖湖南长沙〗
地址：湖南省长沙县暮云西湖村（芙蓉南路
特色：现當放养区内各种凶猛的野生动物。
长沙生态动物园成人票特惠热卖，无等待，快入园
锦绣湘蘸渫卡优待票（几引学生/老人）（问邮〕
长沙世之窗成人票（十一特惠，散拼票
0
长沙世界之窗
〖湖南长沙〗
地址：湖南省长沙市开福区三一大道485号
特色：现全球标志建筑及异国风情表濱。
长沙世之窗成人票
考门市价
¥105
¥200
考门市价
¥170
¥165
4．4分（1454人点评）
查看详情卜
兑换方式
预付
预付
查看吏多（6）-
4．5分（2815人点评
查看详情卜
兑换方式
预付
预付

**2.酒店信息爬取**

使用这条http://hotel.meituan.com/changsha/

从数据库中获取城市名字

然后使用pypinyin库,把文字转化为拼音,调用这条url获取中国所有城市的酒店实时计价

pypinyin教程在

<https://blog.csdn.net/r_coder/article/details/79419318>

非常简单

这样获取酒店和所属城市的信息,完善城市信息

计算机生成了可选文字:
智能排序
价格
0
好评优先人气优先
0五彩今天连锁氵酉店（师范学院高铁站店）
[酃湖大字城]珠区酃骼与衡花骼交氵卜东南尾看地
无线上网@停车场0行李寄存餐厅
．诗月主题客栈（师范学院高铁站店）
[酃湖大字城]珠区酃骼衡阳帅范字院南门对面美街巷内
看地認
无线上网0行李寄存热水洗氵谷电吹风
0后花园时尚酒店
[石鼓区]石鼓区蒸水桥北晶珠广场（中国电信斜对面，华源市场
近）看地
无线上网有线上网@停车场0行李寄存
4．6分很好
5開0十人氵肖费
看评论
4．9分很好
5開0十人氵肖费
看评论
4．2分好
5開0十人氵肖费
看评论
50分钟前有人预订
查看详情
29分钟前有人预订
看详情
¥138起
1小时前有人预订
查看详情