

Python 项目——规划旅行小能手（Travel Web）介绍

Python2 班 第 N 小组

小组成员：何智钧、陈宛佳、林元靖

一、 项目名称

中文名称：规划旅行小能手

英文名称：Travel Web

二、 项目背景

考虑到项目的长远意义和现实意义，我们选择了旅游这个主题。“旅”是旅行、外出，即为了实现某一目的而在空间上从甲地到乙地的行进过程；“游”是外出游览、观光、娱乐，即为达到这些目的所作的旅行——二者合起来即旅游。

作为人们开拓视野、放松身心的方式，旅游的普适性很高，无论小孩还是中老年人，都可以通过不同的旅游方式、旅游景点和旅游路线找到适合自己的旅游模式，因此，旅游这个主题，可以说是永恒的。

但“旅”和“游”并存，要想真正地达到游览、娱乐的目的，路线的规划还是十分重要的，虽然说目前已经存在有许多的旅行社和旅游网站，但人们如果想要自行规划旅游路线并不轻松。

因此，我们希望能够利用一个学期学到的 Python 原理加上从网络上获取的更多知识来实现一个为旅游提供信息、规划路线的网站，以简化自行规划的过程。

三、 解决的问题

我们考虑到许多大旅游网站都是景点、酒店、交通和旅游攻略的信息汇总，主要适用于前期方向的大致规划，但因为信息过多，如果要进行细致的规划还需要花费很长的时间筛选信息。

因此，我们的网站减少了交通和攻略的呈现，添加了未来天气状况的信息和规划路线的功能，主要解决了中期细致规划过程中寻找最佳路线以最大程度节省路途时间的问题，并对后期准备及旅行过程中查看未来几天天气和及时调整路线的问题提供了帮助；同时，我们对物美价廉的旅社信息进行了查找和汇总，方便用户了解住宿信息。

四、 项目亮点

1. 所有模块集合在一个基于 `django` 库网站上进行展示，操作简单，集约性、实用性强。
2. 本网站所有信息都是即时信息，符合网站的形式，同时对于即将出行的用户来说具有宝贵的价值。
3. 在旅社和热门景点的网页中都设置了超链接，方便用户进入相关网站查询更详细的信息或进行购买决策。
4. 注重与用户的交互性。用户自行选择查看全国任何城市的相关内容；用户还可以选择想去的景点，网站为其提供最佳路线的设计（尚未找到获得两地点间最短公里数的方法，因此以直线距离代替）；还可以选择查看一定价格区间的旅社信息。
5. 注重信息的总结和数据的统计。无论是热门景点及路线、旅社价格还是天气状况，我们都用科学的方式进行统计，并采用 `pygal` 和 `pyecharts` 等库设计好看的图表或是地图进行展示。

五、 实现网站功能的重要类和方法

在网页模型 `models.py` 中内容的代码主要由 9 个重要的类组成：

1. **RememberCities**：将用户搜索过的城市及其简介信息放入 `csv`
2. **RetellCities**：按用户搜索的频率由高到低返回城市及其简介信息
3. **CityInfo**：从百度百科爬取城市简介，剔除多余字符后返回
4. **HotDestination**：从马蜂窝爬取城市的热门城市，对热度前十五位的景点进行柱状图展示；从百度地图开放平台爬取景点的经纬度，并在城市地图上呈现所有景点的分布。
5. **WeatherForecast**：从百度地图开放平台爬取城市天气预报信息，并进行折线图和图表展示。
6. **HostelPrice**：从木鸟民宿网站爬取城市旅社信息，利用饼图统计价格信息并利用图表展示对应价格区间旅社的基本信息。
7. **HostelIntroduction**：用户选定价格区间后筛选旅社信息，并用表格和地图

形式呈现。

8. **BestRoute**: 根据用户选择的景点，为其设计最短旅行路线，并在地图上进行展示。

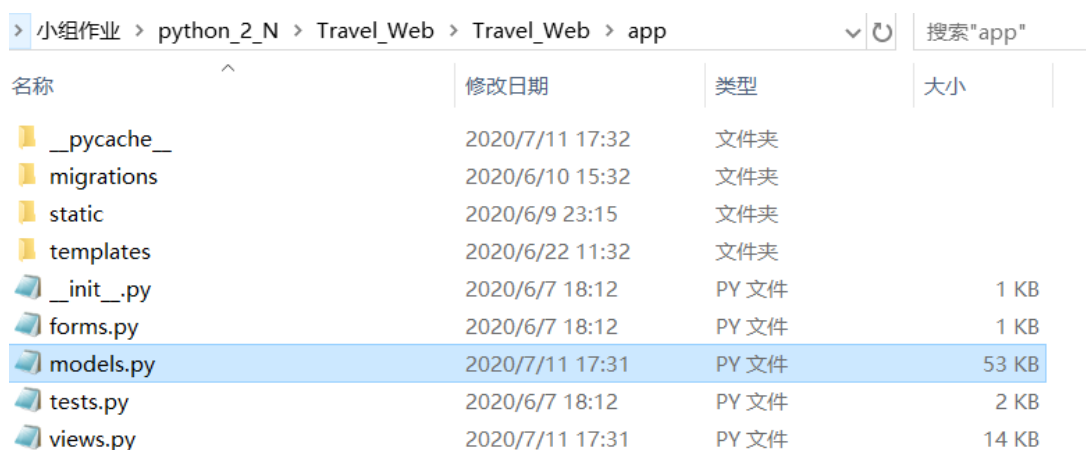
9. **RecommendRoute**: 从途牛网站获取某一特定城市的推荐旅行线路。

六、 代码分布

打开 Travel_Web 文件夹，里面的 Travel_Web.sln 是 Visual Studio 的项目文件，若安装了 Visual Studio 则可以直接打开并运行。

打开内一层的 Travel_Web 文件夹，manage.py 是 django 运行整个网站的文件（未作修改）。

1. 打开内二层的 Travel_Web 文件夹，settings.py 是 django 网站的设置文件（未作修改）；urls.py 定义了网站中具体包含哪些网址（共 11 个），当用户发出请求的时候，urls.py 把用户请求的网址匹配到网页视图。（视图是什么呢？下面会提到。）
2. 打开内二层的 app 文件夹，models.py 的占用空间最大（如图 1），里面把整个网站需要用到的方法（比如说爬虫、统计、作图等等）打包成类和函数，在 views.py 中调用；views.py 就是定义网页视图（共 11 个）的文件了，它接收用户发过来的请求，经过运算之后，把它们呈现在网页上。（呈现的格式是什么样的呢？马上会提到！）



名称	修改日期	类型	大小
__pycache__	2020/7/11 17:32	文件夹	
migrations	2020/6/10 15:32	文件夹	
static	2020/6/9 23:15	文件夹	
templates	2020/6/22 11:32	文件夹	
__init__.py	2020/6/7 18:12	PY 文件	1 KB
forms.py	2020/6/7 18:12	PY 文件	1 KB
models.py	2020/7/11 17:31	PY 文件	53 KB
tests.py	2020/6/7 18:12	PY 文件	2 KB
views.py	2020/7/11 17:31	PY 文件	14 KB

图 1 model.py 与 views.py

3. 打开内三层的 template 文件夹，打开内四层的 app 文件夹，里面有很多 html 文件（如下页图 2），其实它们定义了每个网页的格式，可以用记

事本打开，它们的语法形式虽然与 py 文件大不相同，但基本逻辑是一致的。需要注意的是双层大括号里面的东西是变量名；也有形式类似的条件语句和循环语句。

python_2_N > Travel_Web > Travel_Web > app > templates > app					搜索"app"
名称	修改日期	类型	大小		
best_route.html	2020/7/11 17:24	Firefox HTML D...	1 KB		
city_introduction.html	2020/7/11 17:21	Firefox HTML D...	2 KB		
hostel_introduction.html	2020/7/11 17:23	Firefox HTML D...	2 KB		
hostel_price.html	2020/7/11 17:22	Firefox HTML D...	1 KB		
hot_destinations.html	2020/7/11 17:21	Firefox HTML D...	5 KB		
my_about.html	2020/7/11 17:19	Firefox HTML D...	4 KB		
my_contact.html	2020/7/11 17:18	Firefox HTML D...	1 KB		
my_index.html	2020/7/11 17:17	Firefox HTML D...	2 KB		
my_layout.html	2020/7/11 17:31	Firefox HTML D...	3 KB		
recommend_route.html	2020/7/11 17:35	Firefox HTML D...	2 KB		
search.html	2020/7/11 17:19	Firefox HTML D...	3 KB		
search_result_nothing.html	2020/7/11 17:25	Firefox HTML D...	1 KB		
weather_forecast.html	2020/7/11 17:22	Firefox HTML D...	2 KB		

图 2 html 格式文件

- 打开内三层的 template 文件夹,打开内四层的 data 文件夹,city_pinyin.csv 和 weather_city_id.csv 是把城市名分别和拼音与代码对应; history_search.csv 用来记录用户历史搜索的城市; 其它几个 svg 和 html 文件是临时图片和网页文件,每次生成对应的网页都会临时储存在这里,需要呈现的时候就从这里提取。
- 打开内三层的 static 文件夹,打开内四层的 app 文件夹,里面是网站的样式、字体等文件。

七、运行方式及解说

- 需要安装最新版本的 django、pyecharts、pygal 等库。
- 推荐使用火狐和谷歌浏览器,使用 120%的视图;不建议使用 Edge。
- 如果安装了 Visual Studio 则可以直接在项目文件中直接运行,生成网站;如果不是使用 Visual Studio 的话可以参照课本 359 页,在 manage.py 文件夹下用命令提示符输入 python manage.py runserver,然后复制网址到浏览器打开即可(在使用过程中请勿关闭命令提示符窗口)。
- 考虑到与自带字体的匹配性,本网站的格式语言统一使用英语。

5. 呈现在眼前的是 Travel Web 的首页（如图 3）。网页上方的黑色框和按钮以及下方的时间写在了网页格式的父模板里，每个网页格式的子模板都继承了这个父模板，所以，所有网页都会出现上方的黑色框和下方的时间（是发出请求进入这个页面的时间）。网页的名称写在浏览器的标签栏里。上方的 Travel Web 和 Home 指向的都是首页的网址；About 链接到网站的介绍；Contact 链接到我们小组成员的联系方式；Search 链接到搜索城市页面，这样无论在网站的哪个页面，都有通过这个按钮开始城市的搜索。

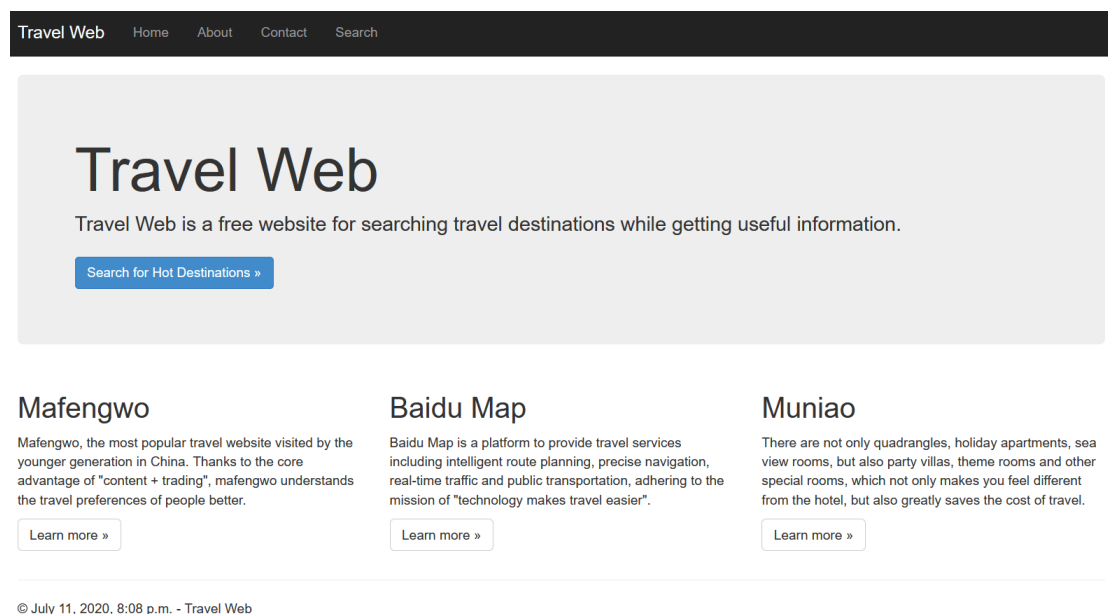


图 3 首页

6. 再来看看首页的其它元素。中央的 Search for Hot Destinations 和上方的 Search 链接到同样的搜索页面；下方是 3 个友情链接，点击 Learn More 会链接到它们的官方网站。
7. 现在进入搜索页面（如下页图 4）。左上方是选择城市的搜索框，用户在其中选择一个城市，点击 Search 即可；这个搜索框不允许用户乱输入，比如输入“哈哈”，它是不会提交任何信息的，然后会跳转到 Search Nothing。下方是 3 个推荐搜索的城市及其介绍的前 250 个字，它们是用户历史搜索次数最多的城市；如果用户是初次使用或者搜索过的城市不足 3 个，则依次添加北京、上海、广州及其介绍。（除此以外，网站的其它地方都是使用即时数据。）实现方法大致是用对历史搜索进行排序，把字典列表传输到视图，然后在 html 文件中选择排在最前面的三个进行呈现。

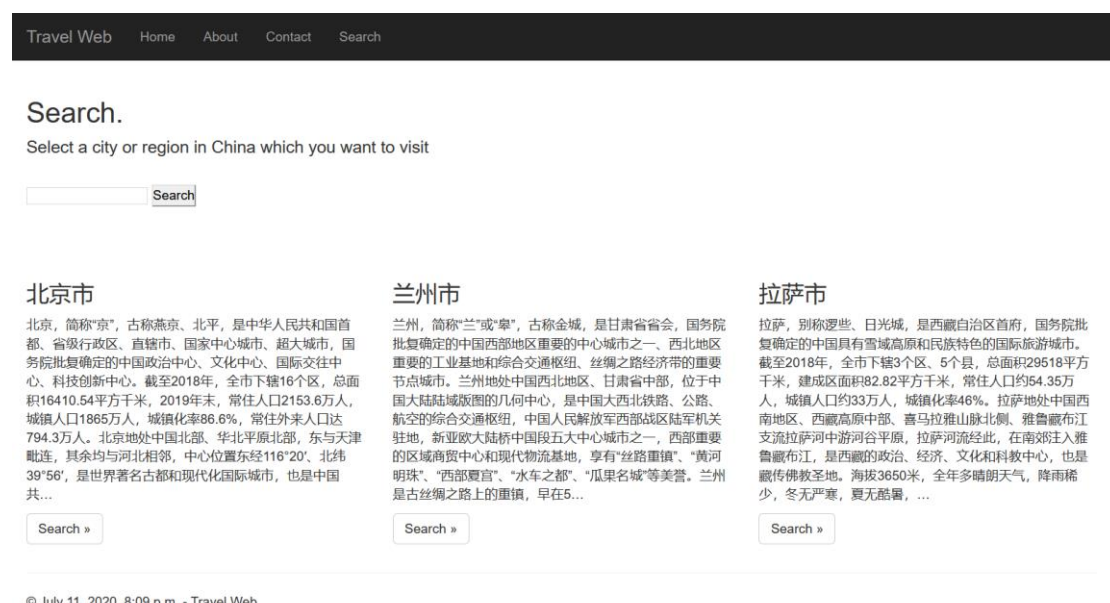


图 4 搜索页

8. 选择一个城市并点击 **Search**，就会见到这个城市的介绍。实现方法是从百度百科提取城市简介并返回给网页视图；与此同时，记录历史搜索的 **csv** 文件就会写入这个城市及其介绍，用于下一次的推荐搜索。下方三个链接分别链接到这个城市的天气、景点和旅社，可以依次点开。
9. 在天气预报网页（如下页图 5），首先看到的是这个城市未来 5 天的天气情况的表格，下面是未来 5 天的最高温度和最低温度。鼠标指向图上每个点就会显示对应的温度。这个 **XY** 散点图用埃尔米特插值法进行连线，所以曲线比较光滑。实现方法是从百度地图获取天气预报，其中最高和最低气温列表用于作图，白天和晚上的天气列表就直接返回给网页；网页中用表格形式呈现天气，并从文件夹提取刚刚保存的气温变化图。注意，若不能搜索到天气情况，则表格显示 **N.A**，统计图显示 **No Data**。

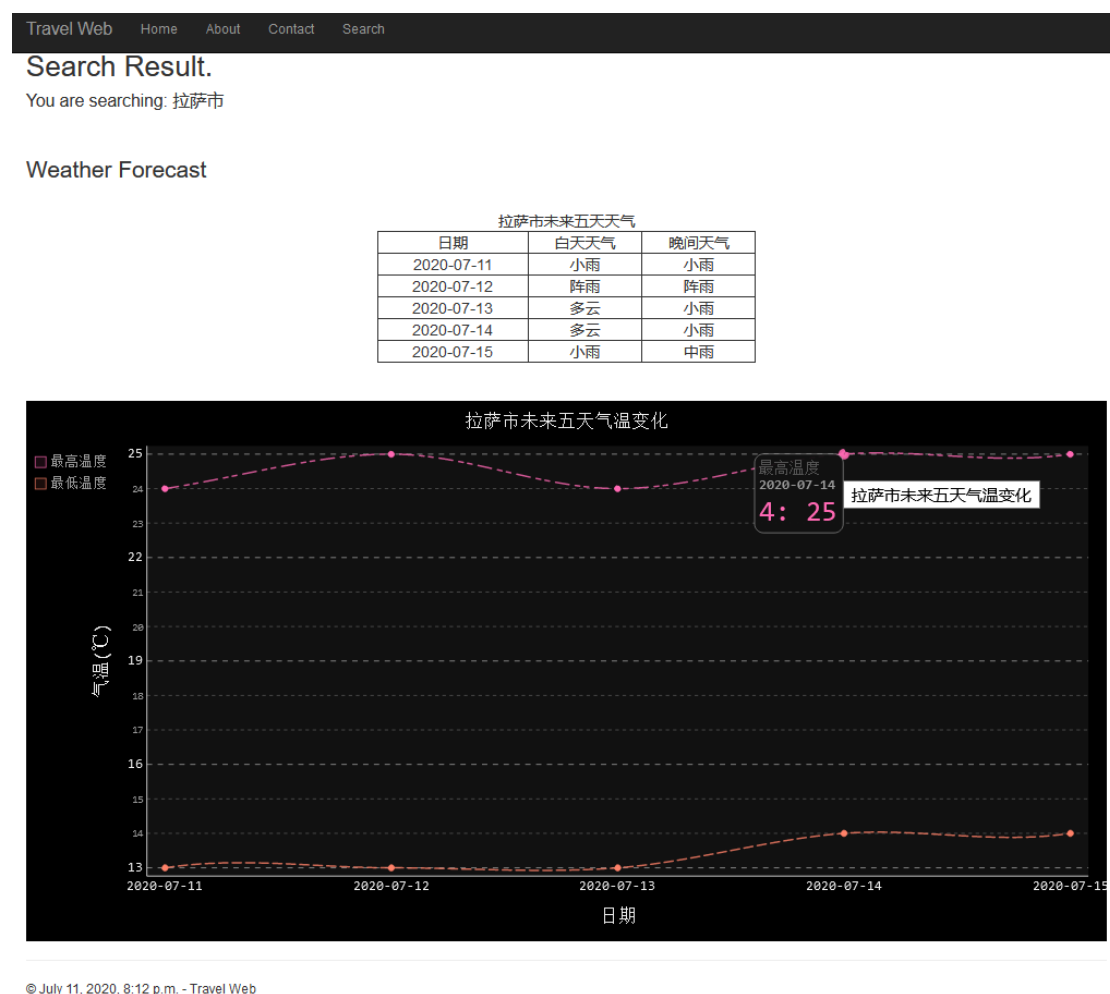


图 5 天气预报页

10. 跳转到热门景点页面时间相对较长。在热门景点网页，首先可以看到这个城市最热门的景点以及它们在马蜂窝上的点评数量和游记数量(如下页图 6)。我们这几张统计图都是采用“霓虹灯”风格，动作渐变特效比较多。横坐标是景点的名称，如果多于 5 个字会省略，当鼠标移动至其上方时，则会显示全名。把鼠标移动至柱状图的上方会显示对应的点评和游记数量；同时我们可以发现，这里是带了链接的——链接到什么呢？把鼠标指向标题，原来是链接到景点的地址。无论是点击柱状图还是点击提示中的 Link，都会跳转到百度地图页面。实现方法是从马蜂窝获取景点的点评和游记数量，并根据点评数量进行排序（我们设定了呈现前 15 个景点，如果搜索到 $n < 15$ 个则呈现 n 个），作图的时候附上百度地图链接；和刚才天气的统计图相似，在网页呈现的时候从特定位置提取热门景点统计图。若不能搜索到景点，则统计图显示 No Data。

Travel Web Home About Contact Search

Search Result.

You are searching: 拉萨市

Hot Destinations & Links

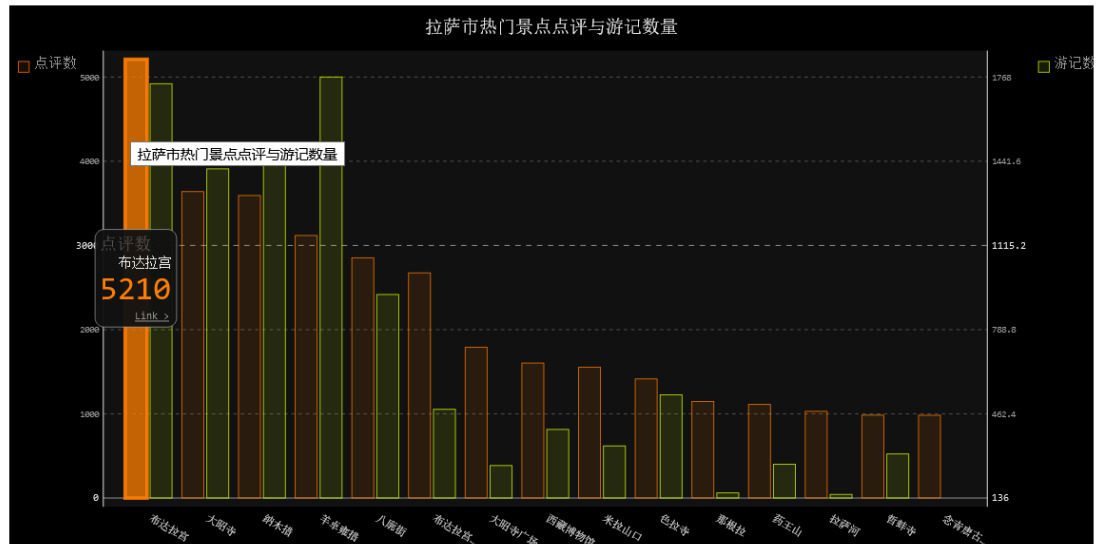


图 6 热门景点统计图

11. 统计图下方是搜索到的景点在地图上的呈现（使用涟漪特效，如图 7），由于点评数量和游记数量差别较大，我们使用加权平均的方法定义热度指数，

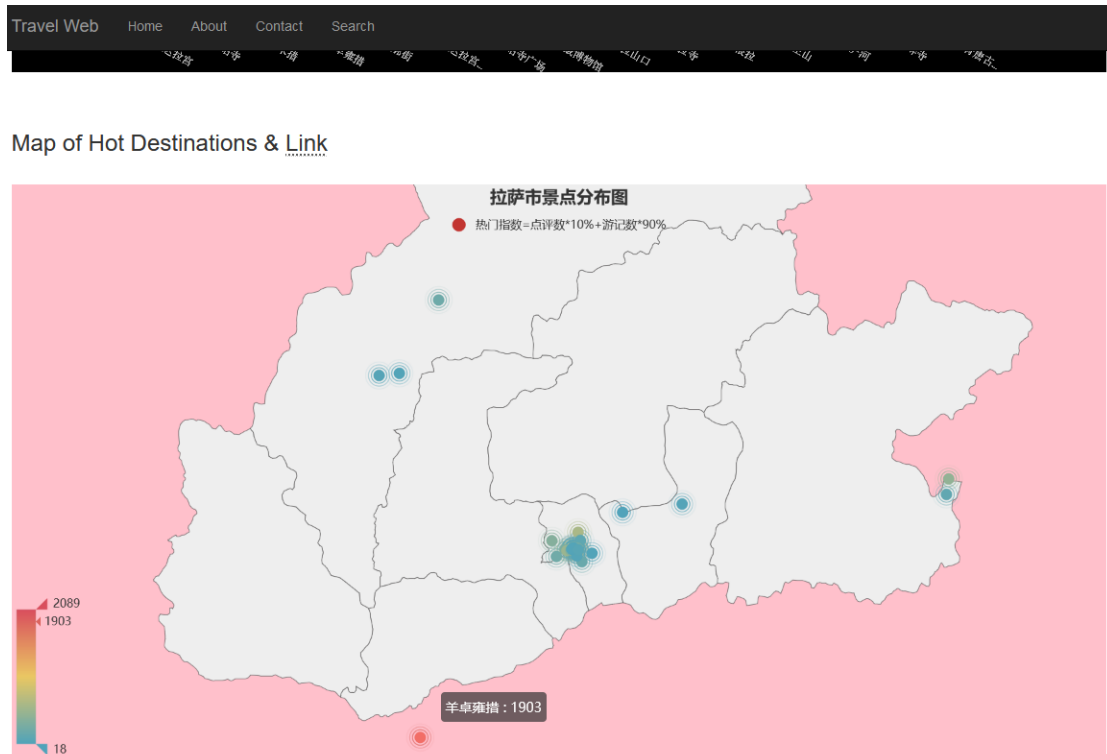


图 7 景点分布图

同时对统计图中最大、最小值进行定义。地图的标题链接到该城市在马蜂窝上的相应页面。实现方法是获取该城市中这些景点的经纬度，画出地图并保存在特定文件夹内，在网页呈现的时候从特定位置提取。

12. 地图的下方是推荐搜索的下拉列表框，用户可以在下拉列表框中选择旅行天数（不选择则搜索全部天数对应的路线），跳转到的新页面呈现旅行路线的表格。需要注意的是，旅行路线来自途牛网站，请仔细甄别。
13. 推荐路线下拉列表框的下面则是可供用户选择想去最多 8 个（如果超过 8 个，则运行时间过长）景点的复选框，用户还可以指定起点和终点，我们给用户画出一条最短路径。在网页中，当能检索到该城市的景点的时候，我们使用 `for` 循环制作起点终点的下拉列表框和景点复选框。
14. 点击最佳路径可以跳转到最佳路径页面（如图 8）。这个地图和刚才页面的地图十分相似，不同之处在于用动态箭头画出了用户选择景点的最佳路线。实现方法也比较类似：在马蜂窝获取景点及其热度；获取经纬度；根据用户输入情况寻找最短路径；最后在地图上呈现所有景点以及最佳路线连线。或许要问，为什么要重复搜索呢？认真观察可以发现，我们的网址是有一定规律的，心急的用户可能会直接输入网址，这样我们的网站也能做出正确应答。（本网站每一个网页都是独立的！）或许还要问，为什么网址是先景点后城市呢？其实这种形式的网址是一个字典，键不可以重复，所以我们为了同时传输景点和城市，只能这样反过来设计复选框的 `name` 和 `value` 了。

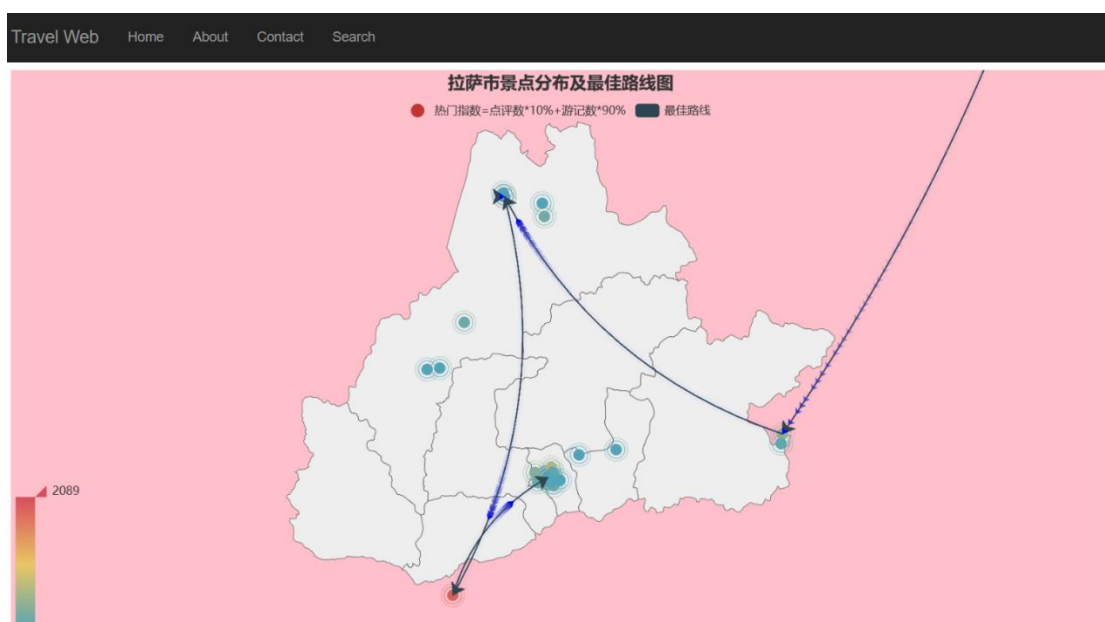


图 8 最佳路线图

15. 跳转到旅社价格页面时间相对较长。我们的网站搜索的是旅社而不是酒店信息，所以使用了 `hostels` 一词。首先在木鸟官网遍历对应城市的每一个页面，获取价格信息，然后数一数每个价格区间段有多少个旅社，画出扇形图（如图 9）。鼠标指向每个扇形都会看到相应价格区间的旅社数量；点击链接就会跳转到相应旅社的具体介绍（在这张图中，扇形图和标签都加了链接）。

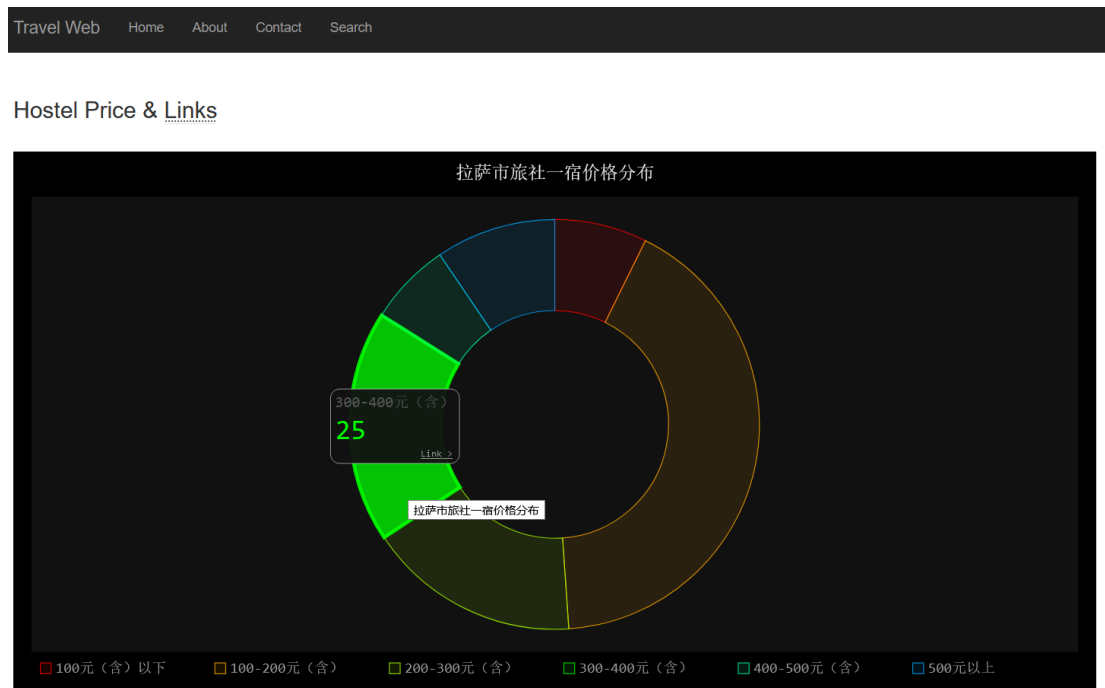


图 9 旅社价格统计图

16. 跳转到旅社介绍页面时间相对较长。与刚才略有不同的是，这次遍历木鸟官网页面的时候把名称、地址、基本信息等和价格一并统计了，最后需要返回字典列表的时候根据价格区间做出筛选。在网页中，使用 `for` 循环生成旅社介绍表格；如果搜不到数据，也就是列表为空的时候显示 N.A。在表格最后一列网址部分加入了链接按钮，点击它可以进入木鸟中的介绍页面。下方则是这些酒店的地图呈现了（与热门景点地图类似，在此不再一一赘述，如图 10）。

Travel Web	Home	About	Contact	Search
罗布林卡附近, 藏式风格, 可以自己做饭	1室 整租 宜住2人	138	金珠西路24号	Learn more »
拉萨布达拉宫那里民宿粉红温馨浪漫标准间103 (公卫)	2室 单间出租 宜住2人	108	雪新村宗角新村460号警务站对面	Learn more »
拉萨布达拉宫那里民宿蓝色阳光大床房208 (独卫)	1室 整租 宜住2人	150	幸福社区宗角新村	Learn more »
毗邻八廓街东缘温馨单人房	1室 整租 宜住1人	139	嘎玛贡桑街道	Learn more »
毗邻八廓街东缘豪华大床房	1室 整租 宜住2人	188	嘎玛贡桑街道	Learn more »
统建社区精致标准间	1室 整租 宜住2人	168	嘎玛贡桑街道	Learn more »

Map of Hostels (Price: ¥ 100-200)



© July 11, 2020, 8:27 p.m. - Travel Web

图 10 旅社信息及分布图

八、与用户的交互

1. 网站会记录用户搜索的记录, 在网站首页展示用户搜索频率最高的三个城市及其简介信息。
2. 用户可以选择全国任何一座城市, 查询到相关的旅行信息。
3. 用户可以通过超链接, 快速地去到相关网站进行更详细的信息查询。
4. 用户可以根据景点分布地图选择多个想要前往的景点, 并得到连接所选择景点的最短路径规划。
5. 用户可以选择特定的价格区间, 了解该价格区间内旅社的信息。

八、项目反思与展望

1. 加入 login in, 为每一个用户记录历史搜索。
2. 向用户展示每一个景点的具体信息。

3. 加入交通方式的模块，为用户提供不同交通方式的信息，方便其进行比较。
4. 使用坐标点之间的直线距离计算最短路径，会存在和实际相差较大的情况。
5. 路径设计结合具体道路和交通的拥挤情况的考虑。
6. 将分散在各个类当中的相似性较高的函数，比如画地图函数、爬取经纬函数等，存放起来进行反复灵活的调用，使代码更简洁。

九、 小组分工

何智钧：代码分类与整合；网站的设计、搭建与内容的展示；柱形图、饼图、折线图等图的制作与美化。

陈宛佳：爬取热门景点信息、推荐路线信息；计算最短路径；制作展示 PPT。

林元靖：爬取城市简介、天气信息和旅社信息；景点、最佳路线、旅社分布的地图可视化。