



计 算 机 网 络 课 程 设 计

兰州招聘岗位的分析

Hollow Man

目录

一、项目简介3

二、技术细节3

 2.1、招聘门户网站的选取3

 2.2、数据爬取4

 2.3、数据清洗7

三、总结12

一、项目简介



近年来，随着经济下行压力的加大和各行各业的发展，大学生毕业后找不到工作也逐渐成为一种新的态势。所以，作为现在的大学生，我们应该如何做好职业规划，选择自己的专业方向，从而适应现在的职场发展，让自己能够在未来的求职大战中胜出呢？



为了解决这个问题，我决定采用现在热门的大数据分析技术，使用 Python 爬虫爬取招聘网站有关兰州的求职信息，随后进行数据清洗，做成各种维度的数据分析，从而能够从中一窥当前兰州市各行各业最紧缺的人才，为求职者提供参考。

二、技术细节

2.1、招聘门户网站的选取

一开始我选取了智联招聘 (<https://www.zhaopin.com/>) 和 boss 直聘

(<https://www.zhipin.com/>) 这两个国内知名的招聘网站进行数据的爬取。但是，在实际操作中发现，boss 直聘对于爬虫非常不友好，当爬取过多数据时，会导致报 403 错误，服务器拒绝地址请求，爬取到的页面提示：“由于您当前网络访问页面过于频繁，可能存在安全风险，我们暂时阻止了您的本次访问，24 小时将自动解除限制。”这种情况需要挂一个新的 ip 代理才能解决，而智联招聘则没有这种问题，因而我最终选择了智联招聘网站作为数据来源。

2.2、数据爬取

我们平时浏览网页时，在浏览器的地址栏中输入一个 URL，浏览器会向 HTTP 服务器发送 HTTP 请求。HTTP 请求方法主要分为“Get”和“Post”两种方法。

当我们向浏览器发送一个 Request 请求去获取一个站点的 html 文件，服务器把 Response 文件对象发送回给浏览器。浏览器分析 Response 中的 HTML，发现其中引用了很多其他文件，比如 Images 文件，CSS 文件，JS 文件。浏览器会自动再次发送 Request 去获取图片，CSS 文件，或者 JS 文件。

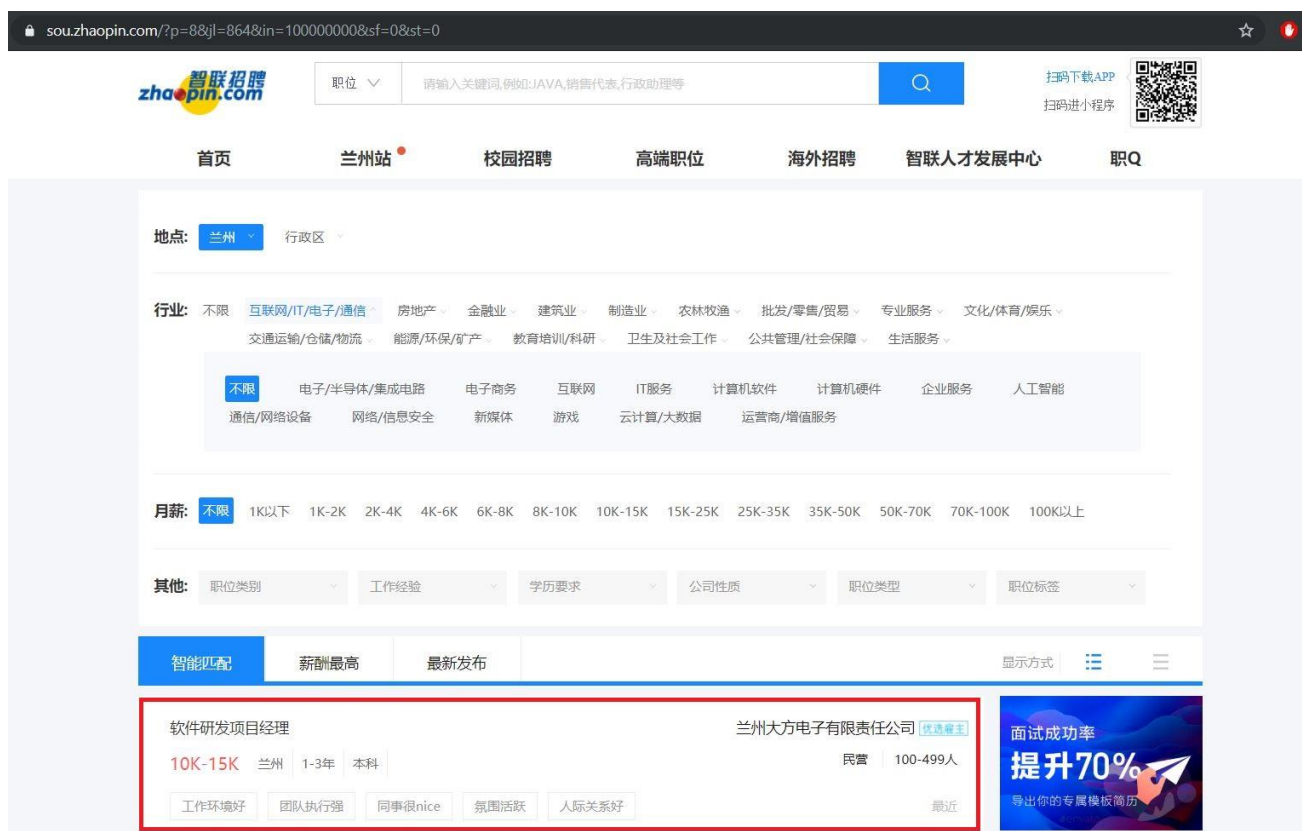
所有的文件都下载成功后，网页会根据 HTML 语法结构，完整的显示出来我们平时通过浏览器访问的页面。

爬虫是面向特定目的需求的一种网络信息爬取程序，其原理就是计算机程序模仿我们普通用户正常的浏览器操作，并且将其获取到的信息进行保存，便于进行进一步的操作。

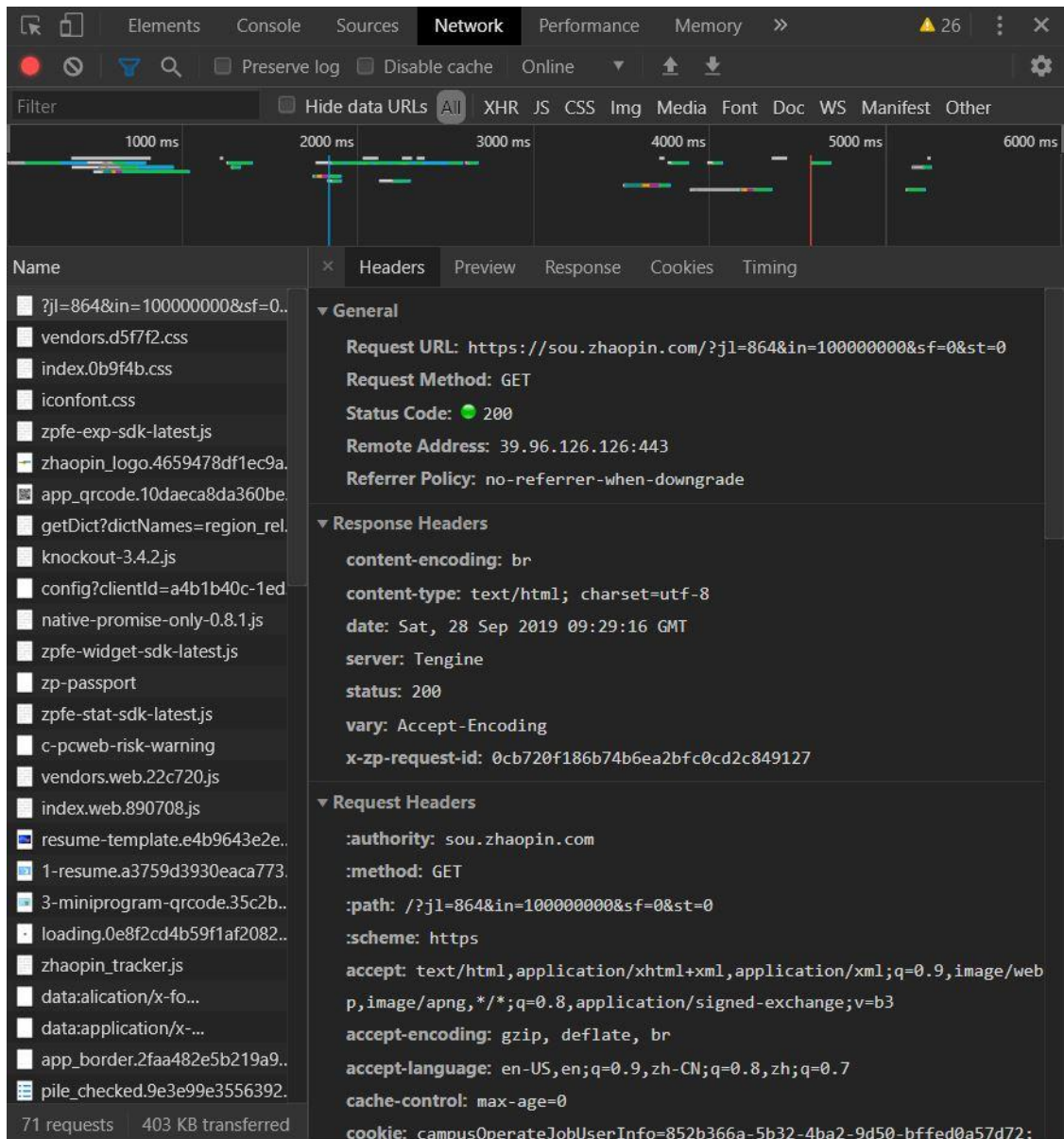
要实现爬虫模拟用户浏览，我们必须准确记录下浏览器发出 Request 请求时的 header 和 cookies 信息，因为现在很多网站都有防爬虫机制，只有将爬虫伪装成用户浏览器的访问，才能实现爬取。

让我们看一看智联招聘的网站，首先打开我们需要爬取的页面，我们需要爬取的是兰州计算机相关岗位的招聘信息。记录下网址，点击下一页时，发现“p=*”

这个参数加一，并且有关兰州计算机招聘的信息总共有 12 页，所以 p 参数为页面参数。同理，通过这种方法，我发现了 jl 是地域参数，in 为职位参数。



接下来，通过谷歌 Chrome 浏览器 F12 开发者模式，我们可以观察到请求的 header 和 cookies 信息，我们需要记录下这些信息。



但是在第一次尝试时，我直接用爬虫进行爬取，发现页面中并未显示任何职位的信息。进一步用谷歌 Chrome 浏览器 F12 开发者模式，我们得到了一点启发。原来数据招聘岗位的有关数据都通过 <https://fe-api.zhaopin.com/c/i/sou?pageSize=90&cityId=864&industry=100000000&salary=0,0&workExperience=-1&education=-1&companyType=-1&employmentType=-1&jobWelfareTag=-1&kt=3> 这个 api 接口，通过这个接口我们便能进行数据的获取。

这边对部分参数做下解释，其中 cityId 是城市编号，其中 864 对应兰州，

start 是起始记录，第一页为 0，第 n 页为 60*(n-1),其中 60 为 pageSize,还有就是 lastUrlQuery 参数，值为 map 类型（python 中的字典、js 中的 json），即包装键值对。

查看响应 response 发现，我们获取的不是 html 字符串，而是一个 json 格式的数据。这个不奇怪，如今使用 Ajax 异步加载服务端数据并渲染数据到页面上已经是很普遍了，这样获取需求数据更加方便，省去了 parse HTML 页面的步骤。

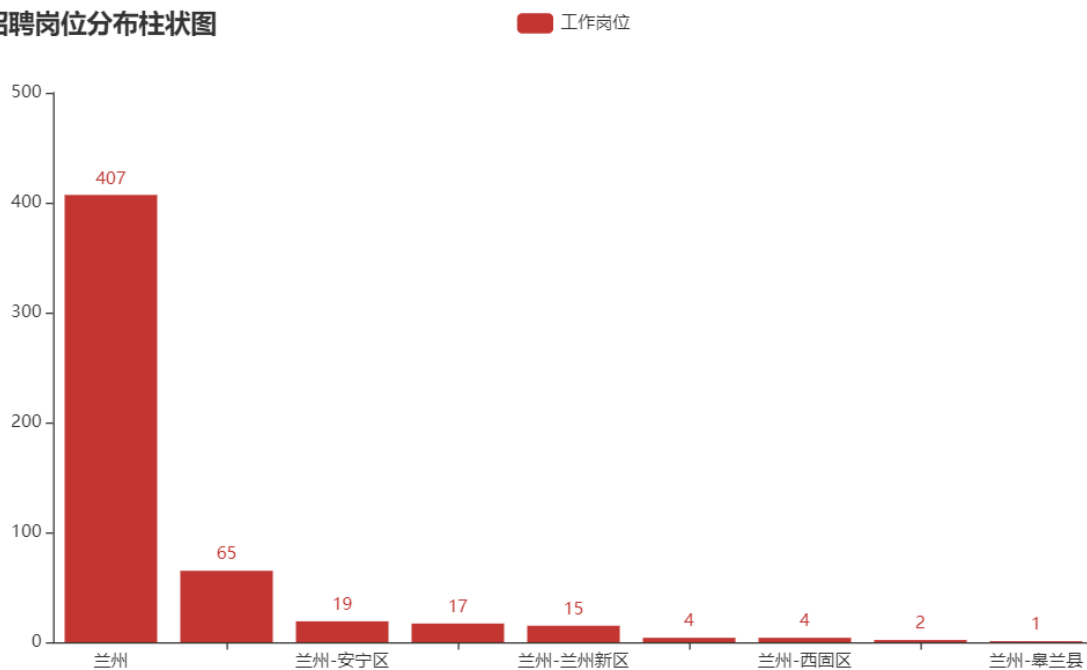
接下来我们就可以开始进行爬虫的编写了，我们使用 requests 库进行数据的获取，并且使用 sqlite 保存我们的数据。

number	jobType	companyName	company_size	company_type	work_years	edu_level	salary	empl_size	jobName	city	welfare
兰州_CC324405513390250296000	市场	兰州立博生物科技有限公司	100-499人	民营	1-3年	中专	4K-6K	全职	市场督导	兰州	
兰州_CC479660631300233994508	市场	甘肃中德文化传媒有限公司	20-99人	其它	不限	大专	2K-4K	全职	市场开发	兰州	包吃,包住
兰州_CC650934121300282186206	商超/酒店/娱乐管理/服务	甘肃华奥恒信装饰工程有限公司	20-99人	其它	1-3年	中专	4K-6K	全职	导购(销售顾问)	兰州	绩效奖金,年终奖,五险一金
兰州_CC566891123300392686101	软件/互联网开发/系统集成	兰州深蓝图形技术有限公司	20-99人	民营	1-3年	本科	4K-8K	全职	java工程师招聘	兰州-城关区	年底双薪,五险一金
兰州_CC673050722300465200901	实习生/培训生/储备干部	新疆和润房地产开发有限公司昌吉分公司	1000-9999人	民营	无经验	本科	4.6K-8K	校园	应届生储备干部(2020届校园招聘)	兰州-城关区	五险一金,年终奖
兰州_CC327851310300185977710	销售业务	新疆通信信息股份有限公司	1000-9999人	股份制企业	不限	本科	7K-9K	全职	渠道支持经理	兰州-城关区	弹性工作,五险一金
兰州_CC35731862400272211207	房地产开发/经纪/中介	碧桂园品牌营销中心-兰兰区域	100-499人	民营	1-3年	大专	薪资面议	全职	置业顾问/高级置业顾问	兰州	五险一金,年终奖
兰州_CC269905180300283968902	影视/媒体/出版印刷	兰州胜博教育咨询有限公司	20-99人	民营	1-3年	大专	4K-6K	全职	影视制作总监	兰州-安宁区	绩效奖金,五险一金
兰州_CC875748330300299970302	人力资源	甘肃心相荣商务服务有限公司	20-99人	民营	3-5年	本科	6K-8K	全职	跟单集团人力资源部经理	兰州	五险一金,节日福利
兰州_CC682982526390250004000	销售业务	甘肃鑫和热力科技有限公司	20-99人	民营	1年...	大专	2K-4K	全职	业务员	兰州	五险一金,节日福利
兰州_CC543507022300411057407	人力资源	兰州宏祥房产经纪有限公司	1000-9999人	民营	1-3年	大专	3K-4.5K	全职	薪酬绩效专员	兰州	节日福利,五险一金
兰州_CC334338030300297376104	其他	兰州博瑞管家家政服务中心	20-99人	其它	无经验	大专	4.5K-6K	全职	课程顾问	兰州	绩效奖金,五险一金
兰州_CC316535019300326641506	销售管理	兰州华联综合超市有限公司	1000-9999人	国企	1-3年	中专	3.5K-4.5K	全职	楼面主管(生鲜/食品/百货)	兰州	年底双薪,五险一金
兰州_CC218628588300381520303	销售管理	兰州红星美凯龙世博家居广场有限公司	100-499人	民营	1-3年	大专	4K-6K	全职	大客户经理	兰州	五险一金,年终奖
兰州_CC667447228300129262903	销售业务	兰州市城关区采兰图德服务中心	20-99人	上市公司	不限	大专	4K-6K	全职	电话销售,电话客服/百合网兰州	兰州	年终奖,五险一金
兰州_CC144109626300413503203	烹饪/料理/食品研发	陕西西金餐饮文化有限公司	1000-9999人	民营	1-3年	学历不限	4K-6K	全职	主理、风味技工	兰州	包吃,包住
兰州_CC144109626300413503203	烹饪/料理/食品研发	陕西西金餐饮文化有限公司	1000-9999人	民营	不限	学历不限	2K-4K	全职	帮厨、帮工	兰州	包吃,包住
兰州_CC515930337990250022000	财务/审计/税务	兰州陇华实业集团有限公司	100-499人	民营	3-5年	本科	5K-8K	全职	财务总监	兰州-城关区	餐补,带薪年假,五险一金
兰州_CC1388782183003350502006	销售管理	四川金鹰集团有限公司	100-499人	民营	不限	学历不限	6K-8K	全职	兰州市场区域经理	兰州	包吃,包住
兰州_CC144109626300413503203	烹饪/料理/食品研发	陕西西金餐饮文化有限公司	1000-9999人	民营	不限	学历不限	2K-4K	全职	运营员	兰州	五险一金,年终奖
兰州_CC566891123300392686101	软件/互联网开发/系统集成	兰州深蓝图形技术有限公司	20-99人	民营	1-3年	本科	4K-8K	全职	软件研发工程师	兰州-城关区	年底双薪,五险一金
兰州_CC372656419390250184000	销售管理	融美科技(北京)股份有限公司	100-499人	民营	1-3年	大专	4K-8K	全职	区域经理	兰州	五险一金,年终奖
兰州_CC386677972300441615805	销售管理	兰州包豪斯装饰工程有限公司	20-99人	其它	3-5年	学历不限	6K-8K	全职	市场部经理	兰州	五险一金,年终奖
兰州_CC858619300300400153103	销售业务	甘肃大鼎源科技股份有限公司	不限	民营	1-3年	高中	5K-10K	全职	电话销售	兰州	
兰州_CC473621013300270983806	销售业务	阳光集团股份有限公司	10000人以上	上市公司	1-3年	本科	薪资面议	全职	地产渠道专员,电话销售实习生	兰州-城关区	五险一金,年终奖
兰州_CC68797433003027142704	销售管理	兰州市城关区采兰图德培训学校	20人以下	民营	不限	大专	8K-10K	全职	招生负责人	兰州	
兰州_CC667447228300115128303	销售业务	兰州市城关区采兰图德服务中心	20-99人	上市公司	1-3年	大专	4K-6K	全职	销售/客服/百合网兰州分公司	兰州	年终奖,五险一金
兰州_CC661276724000479152401	市场	兰州新锦江杉源洋有限公司	100-499人	民营	3-5年	本科	4K-6K	全职	营销策划	兰州	绩效奖金,五险一金
兰州_CC2449466010300218814802	销售业务	上海中增利教育设备有限公司	20-99人	民营	1-3年	学历不限	6K-12K	全职	业务销售	兰州-城关区	项目奖金,五险一金
兰州_CC316535019300326642806	销售管理	兰州华联综合超市有限公司	1000-9999人	国企	1-3年	中专	3.5K-4.5K	全职	楼面主管(生鲜/食品/百货)	兰州-红古区	年底双薪,五险一金

2.3、数据清洗

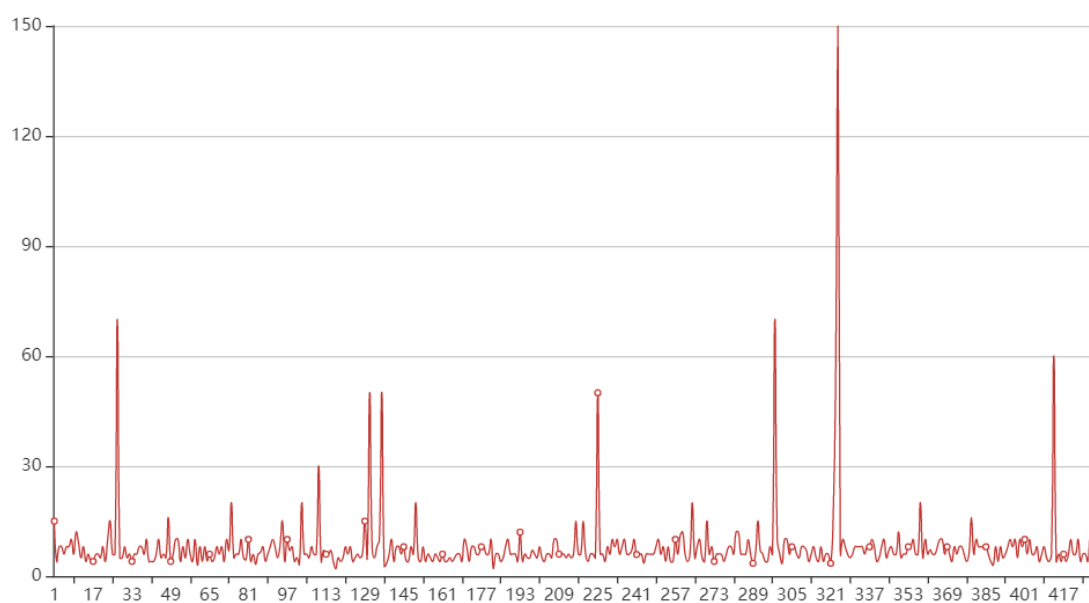
此次爬取共得到兰州招聘岗位 534 个，我们将得到的数据进行数据清洗，并使用 pyecharts 画出数据图表，如附件代码所示。

兰州招聘岗位分布柱状图

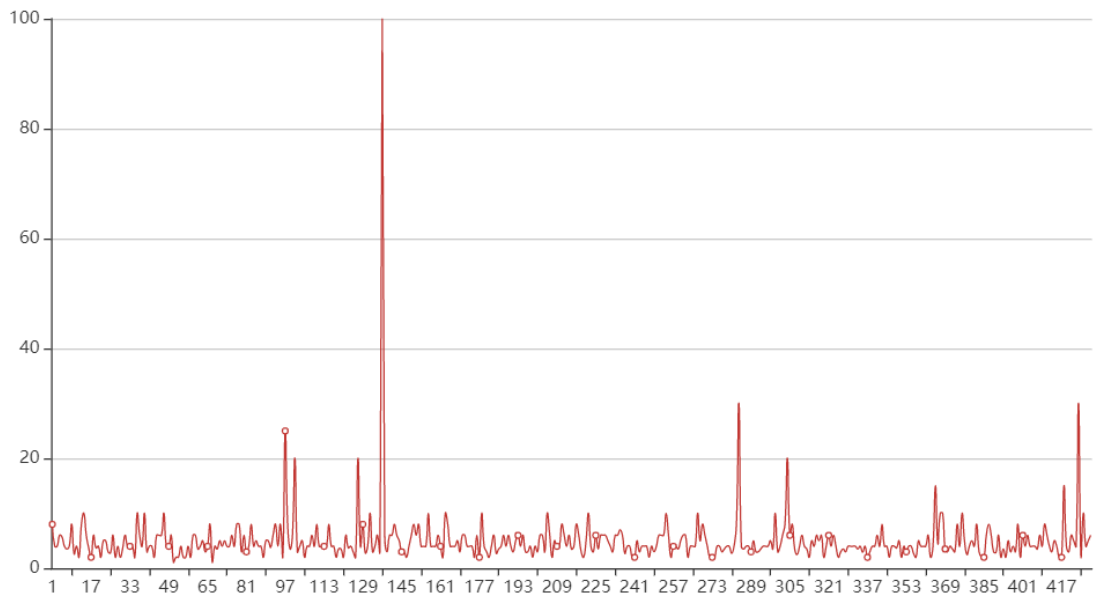


这张图是兰州招聘岗位分布柱状图，我们可以看到兰州城关区人才招聘缺口最大，安宁区，七里河区，新区岗位缺口基本持平。

兰州最高薪水折线图

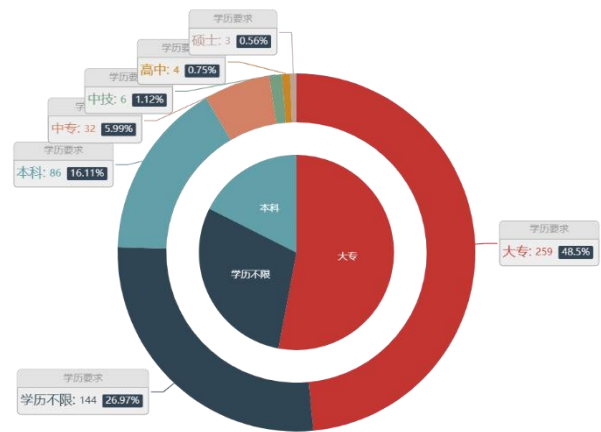


兰州最低薪水折线图



这两张图显示的是兰州最高，最低薪水折线图，由此图可以看出兰州整体工资差别不是很大，除了个别特高薪的工作外，大部分处于平均中间档次。

兰州工作岗位的学历要求饼图



这张图是兰州工作岗位的学历要求，由此图可以看到，兰州市的招聘公司“学历歧视”并非特别严重，对较低学历的求职者而言是个福音。

三、总结

由以上分析，我们可以看到，兰州市的找工作压力并非很大，应届毕业生在兰州市是可以相对容易地找到工作的。同时销售类工作是兰州市最紧缺的工作，因而，以后想要在兰州市工作的同学们可以多多锻炼自己市场营销地能力。

现在刚毕业的大学生都在抱怨找工作难的问题，而为什么经过实际的数据分析，在兰州市找工作似乎并没有这么难呢？这是因为现在大学生都向往着北京上海等一线大城市，不肯到兰州这样的相对较小的二线城市来工作。大城市的工作机会自然会越来越少，相对而言，随着城市化进程和产业转移，在一些小城市，工作机会反而变多了。

因而，作为大学生，我们不应该盲目追求到大城市工作，而应该权衡利弊，选择最适合自己的城市，从而能找到一个相对令人满意的工作。