

甲方名称:

J.D. POWER (君迪)

甲方需求:

- 采集安居客网站全国各城市新房楼盘和二手房小区数据;

- 新房楼盘需要采集的字段如下:

'城市', '楼盘名称', '楼盘单价', '特色', '区域位置', '楼盘地址', '规划总数', '工程进度', '开盘时间', '交房时间', '经度', '纬度', 'URL'

- 二手房小区需要采集的字段如下:

'城市', '小区名', '本月均价', '价格变动', '所在版块', '地址', '总户数', '建造年代', '物业类型', '物业费用', '开发商', '总建面', '容积率', '出租率', '停车位', '绿化率', '经度', '纬度', '简介', 'URL'

- 保证数据完整性, 最终只交付数据;

技术难点:

- 安居客封IP非常厉害, 很容易出现验证码;
- 列表页有页数限制, 不能看到所有页;
- "出租率"是JS动态生成的;

实现方案:

- 使用常规的固定IP代理很容易被封, 必须使用动态IP代理, 每次请求都使用不同的IP;
- 通过"区域"筛选器结合多种排序方式尽可能多的采集数据;
- 使用Python重写JS的生成"出租率"的函数;

城市	楼盘名称	楼盘单价	特色	区域位置	楼盘地址	规划总数	工程进度	开盘时间	交房时间	经度	纬度	
北京	新华联运河	35721	高端	五环至六环-通	运河东岸	1000户	售罄	2012-05-06	2012-10-01	116.686543824	39.925165	
北京	远洋东方公	33295	高端	五环至六环-通	通州区运河西大街运河中学	770户	售罄	2011-07-25	2012-12	116.684793152	39.898151	
北京	中海红杉溪	560	高端	五环至六环-通	通州通燕高速(原京哈高速	556户	售罄	2014-05-25	2015-08-31	116.706425	39.9481	
北京	永定河孔雀城英国馆		飘窗	六环以外-涿州	固安县新昌东街	960户	待售	尚未公开	尚未公开	116.318382	39.463	
北京	润卓天伦湾			飘窗	六环以外-涿州	涿州市华阳东路路北	6473户	待售	尚未公开	尚未公开	116.068443708	39.5103
北京	珠江拉维小	25000	非毛坯	三环至四环-通	永顺镇珠江国际城北侧	1700户	售罄	2010-10-31	2011-09-30	116.69472772	39.942465	
北京	上海8号	55000	非毛坯	三环至四环-朝	朝阳芳园南里9号院	578户	售罄	2011-01-16	2013-06-30	116.491445534	39.976449	
北京	亚奥万和四	59431	非毛坯	五环至六环-朝	北五环顾家庄桥西北角	273户	售罄	2014-08-03	2016-12-30	116.438356	40.030	
北京	中轴国际	80615	非毛坯	二环至三环-东	东城区安定门外西滨河路1	130户	售罄	2009-03-31	2010-07	116.413650864	39.957653	
北京	华彬江南锦		非毛坯	六环以外-昌平	京藏高速28出口(科技园)	20户	售罄	2011-09-21	2013-12	116.145432465	40.205295	
北京	润西山B组	36000	非毛坯	六环以外-门头	门头沟长安街西沿线, 距西	639户	售罄	2014-04-27	2018-09-30	116.105987	39.914	
北京	中海九号公	54545	非毛坯	四环至五环-丰	科丰桥800米路东	280户	售罄	2013-03-29	2013-04-01	116.29709	39.822	
北京	和成锦园	30000	非毛坯	五环至六环-大	大兴东五环大羊坊桥东南角	780户	售罄	2014-04-26	2014-12-31	116.534714013	39.832252	
北京	金成府	31000	非毛坯	五环至六环-通	八通线九棵车站南侧	307户	售罄	2014-03-31	2013-12-31	116.665441	39.895	
北京	富力飞常H	14000	非毛坯	六环以外-大兴	京开高速庞各庄出口南100	444户	售罄	2014-03-01	2015-12-31	116.317699	39.637	
北京	望京金茂府	65407	非毛坯	四环至五环-朝	广顺桥西北500米	92户	售罄	2015-05-09	2015-03-30	116.466396	40.025	
北京	银河湾	60000	非毛坯	三环至四环-朝	西八间房1号	1200户	售罄	2010-05-12	2011-10-17	116.493042	39.987	
北京	贡院9号	53991	非毛坯	二环以内-东	城东城区贡院西街9号	300户	售罄	2009-09	2010-07	116.436894753	39.916587	
北京	西贸国际熙旺中心		非毛坯	六环以外-门头	门头沟双峪环岛	704户	售罄	2009-08-01	2011-06-30	116.119553011	39.941048	
北京	君安国际	31000	非毛坯	五环至六环-大	北京大兴经济技术开发区荣	774户	售罄	2012-07-21	2012-12-31	116.518399	39.795	
北京	观巢	58000	非毛坯	三环至四环-朝	安立路与慧忠北路西北角	116户	售罄	2012-12-30	2013-03-30	116.413740764	40.004953	
北京	富力尚悦居		非毛坯	五环至六环-通	亦庄马驹桥凉水河四支路与	1466户	售罄	2014-03-09	2015-11-01	116.593405331	39.776815	
北京	御墅临枫		非毛坯	四环至五环-海	闵庄路和旱河路东南角	1158户	售罄	2005-11-19	2010-07-30	116.247304352	39.977707	
北京	合生世界花	20000	非毛坯	六环以外-通州	2号桥南500米	2015户	售罄	2013-09-15	2014-05-30	116.553146	39.742	
北京	屯三里	50000	非毛坯	二环至三环-朝	北路与春秀路交汇处	828户	售罄	2009-08-31	2012-05-10	116.451744	39.939	
北京	北京方糖		非毛坯	六环以外-顺义	裕民大街万科城市花园东	100户	售罄	2013-03-01	2014-03-01	116.53670311	40.07491	
北京	国瑞紫金台	90000	非毛坯	二环以内-东	城崇文门外大街(新世界百	323户	售罄	2013-04-30	2013-09-30	116.42592993	39.904648	
北京	逸轩	10500	非毛坯	五环至六环-房	房山长阳镇南营上村	108户	售罄	2014-06-15	2015-09-01	116.223629089	39.770825	
北京	卡法国际		非毛坯	六环以外-顺义	北京市顺义林河开发区, 通	2428户	售罄	1905-07-08	2016-12-30	116.669875943	40.091875	
北京	空港国际	34000	非毛坯	五环至六环-顺	义林河北大街21号	1247户	售罄	2016-04-30	2017-12-30	116.670961745	40.100590	
北京	山水LAVIE	50000	非毛坯	五环至六环-朝	红军营东路	2000户	售罄	2013-01-30	2015-06-30	116.444336368	40.045233	
北京	美然生态馆	25000	非毛坯	四环至五环-大	旧忠路	600户	售罄	2011-12	2011	116.458247182	39.791664	
北京	兴创岭墅		非毛坯	五环至六环-大	后高路与兴丰大街交汇处东	538户	售罄	2013-10-16	2015-12-31	116.343751	39.773	
北京	世茂宫园	54956	非毛坯	二环至三环-朝	东三环北路55号	202户	售罄	2010-11-06	2011-12	116.466496146	39.932308	
北京	京福阁	56194	非毛坯	二环以内-废	弃立交桥东北侧	220户	售罄	2012-10-20	2013-12-28	116.451785	39.905	
北京	京津中央别	9077	非毛坯	六环以外-投	资天津宝坻区	8000户	售罄	2009-02	2012-12	117.316467204	39.72326	

新房楼盘示例数据

城市	小区名	本月均价变动	所在板块	地址	总户数	建造年代	物业类型	物业费	开发商	总建面	容积率	出租率	停车位	绿化率	经度	纬度
鞍山	九街口	2394 比10月下跌	铁西	九街口	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	约50%	暂无数据	0%	(绿化率)122.97027041.1	
鞍山	中大帝景	5913 比10月下跌	铁东	铁东解放东	暂无数据	31-10月-2	暂无数据	暂无数据	1 鞍山市中大	暂无数据	5.49	暂无数据	暂无数据	0%	(绿化率)122.99489641.1	
鞍山	金色年华	4294 比10月下跌	铁东	铁东山南街	暂无数据	20-三月-2	暂无数据	暂无数据	鞍山城投	暂无数据	5.6	约9%	暂无数据	0%	(绿化率)122.98406541.1	
鞍山	新风社区	2978 比10月上涨	铁西	太平村	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	约9%	暂无数据	0%	(绿化率)123.07873141.1	
鞍山	千龙户三期	5154 比10月下跌	铁西	铁西千龙户	暂无数据	暂无数据	暂无数据	暂无数据	鞍山市奥合	暂无数据	暂无数据	约12%	暂无数据	0%	(绿化率)122.98170841.1	
鞍山	福莱特福花	4784 比10月下跌	铁东	福莱特福花	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	约7%	暂无数据	0%	(绿化率)123.03126741.1	
鞍山	健康小学	3414 比10月下跌	铁东	352中四海南	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	约50%	暂无数据	0%	(绿化率)123.02974841.1	
鞍山	嘉联铭座	4915 比10月下跌	铁东	北新兴街	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	0%	(绿化率)123.01371641.1	
鞍山	申江嘉宝花	3945 比10月下跌	铁东	湖南街341F	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	0%	(绿化率)123.02038541.1	
鞍山	地王国际	6168 比10月上涨	铁东	高新区地王	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	约80%	暂无数据	0%	(绿化率)123.06705741.1	
鞍山	马德里皇家	5327 比10月上涨	铁东	立山鞍山胜	暂无数据	暂无数据	暂无数据	暂无数据	鞍山天华	暂无数据	暂无数据	暂无数据	暂无数据	0%	(绿化率)123.04382241.1	
鞍山	恒威滨江国	5056 比10月上涨	铁东	立山胜利北	暂无数据	01-六月-2	暂无数据	1.2	恒威(鞍山)	暂无数据	2.5	暂无数据	暂无数据	0%	(绿化率)123.07575141.1	
鞍山	绿城康庭	6458 比10月上涨	铁西	铁西永乐公	暂无数据	01-六月-2	暂无数据	1	鞍山绿城	暂无数据	2.24	约13%	暂无数据	0%	(绿化率)122.96506241.1	
鞍山	远东嘉山水	7830 比10月下跌	铁东	铁东高新区	暂无数据	0000-00	暂无数据	暂无数据	鞍山市远东	暂无数据	暂无数据	约3%	暂无数据	0%	(绿化率)123.06290141.1	
鞍山	红星美凯龙	3092 比10月上涨	铁西	达道湾	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	约80%	暂无数据	0%	(绿化率)122.93955941.1	
鞍山	永乐朝鲜街	2708 比10月下跌	铁西	永乐朝鲜街	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	约27%	暂无数据	0%	(绿化率)122.96229741.1	
鞍山	韩城公馆	5171 比10月下跌	铁东	韩城公馆	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	约41%	暂无数据	0%	(绿化率)123.02670341.1	
鞍山	立山友好五	3733 比10月下跌	铁东	劳动大厦西	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	约40%	暂无数据	0%	(绿化率)123.02253041.1	
鞍山	深沟寺一区	2625 比10月下跌	铁东	深沟寺一区	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	约33%	暂无数据	0%	(绿化率)123.03144941.1	
鞍山	园丁园	3064 比10月下跌	铁东	海城南台二	660户	0000-00	暂无数据	0.60元/平	海城市南台	115000平方	1.7	约50%	暂无数据	30%	(绿化率)122.79924840.1	
鞍山	马德里花园	5200 比10月上涨	铁东	立山区水源	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	0%	(绿化率)123.04382241.1	
鞍山	菁华家园	4421 比10月上涨	铁东	科技路29号	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	约11%	暂无数据	0%	(绿化率)123.07708341.1	
鞍山	高新区绿色	5555 比10月下跌	铁东	高新区绿色	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	0%	(绿化率)123.05997841.1	
鞍山	大德同德园	3818 比10月下跌	铁东	铁东爱国街	暂无数据	0000-00	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	约66%	暂无数据	0%	(绿化率)123.00491941.1	
鞍山	绿景鑫园	2941 比10月下跌	铁东	海城南台二	660户	0000-00	暂无数据	0.60元/平	海城市南台	115000平方	1.7	约50%	暂无数据	30%	(绿化率)122.79924840.1	
鞍山	前泊国际	6535 比10月上涨	铁东	铁东胜利路	297户	0000-00	暂无数据	1.00元/平	鞍山华达	42000平方	7	约42%	暂无数据	10%	(绿化率)123.00681141.1	
鞍山	新景世纪城	7727 比10月上涨	铁东	铁东新华街	暂无数据	Dec-09	暂无数据	1.20元/平	鞍山新景地	210000平方	6.41	约38%	暂无数据	31%	(绿化率)123.01520641.1	
鞍山	荣誉社区	2972 比10月下跌	铁东	铁东检察院	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	约50%	暂无数据	0%	(绿化率)122.98870941.1	
鞍山	罗马假日	5611 比10月下跌	铁东	罗马假日	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	约4%	暂无数据	0%	(绿化率)123.00085241.1	
鞍山	化里景苑	2222 比10月下跌	铁西	化里景苑	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	暂无数据	约10%	暂无数据	0%	(绿化率)123.00000041.1	

二手房小区示例数据



西安鲲之鹏网络信息技术有限公司

选择我们，所有数据都是你的！

微信扫



公司名称：西安鲲之鹏网络信息技术有限公司

网 址：<http://www.site-digger.com/>

地 址：陕西省西安市雁塔区长安中路99号长安文化综合大厦11821室

邮 编：710061

联系电话：029 - 87553281

手 机：13571845363 齐工

13389148466 周工

客服 QQ：1649677458 或 312602670

邮 箱：hello@site-digger.com