

甲方名称:

某科技股份有限公司

甲方需求:

- 抓取目标网站:
<http://www.digikey.com/>
<http://www.findchips.com/>
<http://www.arrow.com/>
<http://www.verical.com/>
<http://www.futureelectronics.com/>
<https://www.avnetexpress.com/>
<http://www.mouser.com/>
<http://www.chip1stop.com/>
- 抓取数据内容: 定制采集网站上的所有产品数据, 包括但不限于: 每一个产品型号下的分类、品牌、通用参数、细化参数、阶梯价格、技术文档地址、图片地址等数据;

技术难点:

- 数据量大, 采集周期较长, 对代理IP的需求较多;
- 访问频率过高会被网站封IP;
- 产品列表页有页数限制, 不能看到所有页;

实现方案:

- 通过大量稳定高匿HTTP代理IP轮换发出请求, 并严格控制每个IP的两次访问间隔, 以有效防止请求被网站拦截;
- 采用多线程技术, 加快采集进度;
- 通过网站上可用的筛选器, 尽可能多的让列表页页数减少;
- 数据信息量多, 且信息项因产品的不同而变化, 故部分信息项合并保存于 'Attributes' 字段中;

数据示例:

A	B	C	D	E	F	G	H	I
Category1	Category2	Brand	Manufacturer Part Number	Digi-Key Part Number	Stock	Price	Description	Datasheet
2	Audio Products	Accessories	Mallory Sonalet Products Inc.	SG1	458-1043-ND	[0": 0.0]	GASKET FOR SEALING NEMA APPLIC	http://www.mallory-sonalert.com/specifications/S
3	Audio Products	Accessories	Horn Industrial Co LTD	01.AE.MT.601803-G	359-1080-ND	[0": 0.0]	MIC HOLDER FOR 6.0X1.8MM MIC	http://media.digikey.com/pdf/Data_Sheets/Horn In
4	Audio Products	Accessories	Horn Industrial Co LTD	01.AE.MT.976509-G	359-1082-ND	[0": 0.0]	MIC HOLDER FOR 9.7X6.5MM MIC	http://media.digikey.com/pdf/Data_Sheets/Horn In
5	Audio Products	Accessories	Horn Industrial Co LTD	01.AE.MT.401504-G	359-1078-ND	[0": 0.0]	MIC HOLDER FOR 4.0X1.5MM MIC	http://media.digikey.com/pdf/Data_Sheets/Horn In
6	Audio Products	Accessories	Horn Industrial Co LTD	01.AE.MT.974501-G	359-1081-ND	[0": 0.0]	MIC HOLDER FOR 9.7X4.5MM MIC	http://media.digikey.com/pdf/Data_Sheets/Horn In
7	Audio Products	Accessories	Knowles	009P-253159	423-1227-ND	[1": 24.11, *100": 18.07]	PROGRAMMING ACCESSORIES	http://www.knowles.com/eng/Media/Configuration
8	Audio Products	Accessories	Horn Industrial Co LTD	01.AE.MT.946502-G	359-1077-ND	[0": 0.0]	MIC HOLDER FOR 9.4MM MICROPHONE	http://media.digikey.com/pdf/Data_Sheets/Horn In
9	Audio Products	Accessories	Horn Industrial Co LTD	01.AE.MT.601503-G	359-1079-ND	[0": 0.0]	MIC HOLDER FOR 6.0X1.5MM MIC	http://media.digikey.com/pdf/Data_Sheets/Horn In
10	Audio Products	Accessories	Panasonic Electronic Components	M3CA0207	P9952-ND	[0": 0.0]	MIC HOLDER WM-52B/54B/55A D/56A	http://www.panasonic.com/industrial/components
11	Audio Products	Accessories	Knowles	009P-253888	423-1229-ND	3 [1": 25.98, *100": 19.89]	PROGRAMMING ACCESSORIES	http://www.knowles.com/eng/Media/Configuration
12	Audio Products	Accessories	Panavise	105108B	105108BPV-ND	61 [1": 22.99]	SPEAKER STUD MOUNT BLACK	
13	Audio Products	Accessories	Mallory Sonalet Products Inc.	PW1	458-1043-ND	9 [1": 0.32, *100": 0.172]	WASHER COMPRESSION SNP SERIES	http://www.mallory-sonalert.com/specifications/P
14	Audio Products	Accessories	Knowles	BF-1980-000	423-1219-ND	58 [1": 81.38, *10": 78.66]	ACOUSTIC DAMPER WHITE 680OHMS	http://www.knowles.com/eng/content/download/3
15	Audio Products	Accessories	Horn Industrial Co LTD	01.AE.MT.976509	359-1041-ND	[0": 0.0]	MIC HOLDER FOR 9.7X6.5MM MIC	http://media.digikey.com/pdf/Data_Sheets/Horn In
16	Audio Products	Accessories	Panasonic Electronic Components	M3CA0082	P9953-ND	[0": 0.0]	MIC HOLDER FOR WM-34 SERIES	
17	Audio Products	Accessories	Knowles	BFG-30035-G23	BFG-30035-G23-ND	[25": 548.78]	ACOUSTIC DAMPER	
18	Audio Products	Accessories	Knowles	WB-23615-000	WB-23615-000-ND	[0": 0.0]	WAX PROTECTION	
19	Audio Products	Accessories	Horn Industrial Co LTD	01.AE.MT.401504	359-1037-ND	[0": 0.0]	MIC HOLDER FOR 4.0X1.5MM MIC	http://media.digikey.com/pdf/Data_Sheets/Horn In
20	Audio Products	Accessories	Horn Industrial Co LTD	01.AE.MT.601803	359-1039-ND	[0": 0.0]	MIC HOLDER FOR 6.0X1.8MM MIC	http://media.digikey.com/pdf/Data_Sheets/Horn In
21	Audio Products	Accessories	Horn Industrial Co LTD	01.AE.MT.974501	359-1040-ND	[0": 0.0]	MIC HOLDER FOR 9.7X4.5MM MIC	http://media.digikey.com/pdf/Data_Sheets/Horn In
22	Audio Products	Accessories	Knowles	WB-23610-000	WB-23610-000-ND	[0": 0.0]	WAX PROTECTION	
23	Audio Products	Accessories	Horn Industrial Co LTD	01.AE.MT.946502	359-1022-ND	[0": 0.0]	MIC HOLDER FOR 9.4MM MICROPHONE	http://media.digikey.com/pdf/Data_Sheets/Horn In
24	Audio Products	Accessories	Knowles	009P-253747	009P-253747-ND	[50": 921.7]	CABLE PROGRAMMING	http://www.knowles.com/eng/Media/Configuration
25	Audio Products	Accessories	Knowles	2005-253240	2005-253240-ND	[25": 175.45]	SWITCH SLIDE	
26	Audio Products	Accessories	Knowles	BFG-30035-G22	BFG-30035-G22-ND	[25": 548.78]	ACOUSTIC DAMPER	
27	Audio Products	Accessories	Knowles	009P-253748	009P-253748-ND	[50": 921.7]	CABLE PROGRAMMING	http://www.knowles.com/eng/Media/Configuration
28	Audio Products	Accessories	Knowles	2800-253377	2800-253377-ND	[25": 85.91]	SWITCH PUSHBUTTON	http://www.knowles.com/eng/Media/Configuration
29	Audio Products	Accessories	Horn Industrial Co LTD	01.AE.MT.601503	359-1038-ND	[0": 0.0]	MIC HOLDER FOR 6.0X1.5MM MIC	http://media.digikey.com/pdf/Data_Sheets/Horn In
30	Audio Products	Accessories	Knowles	2800-253947	2800-253947-ND	[25": 84.46]	SWITCH PUSHBUTTON	http://www.knowles.com/eng/Media/Configuration
31	Audio Products	Accessories	Switchcraft Inc.	SC600	SC600-ND	[1": 52.5]	DUAL ADAPTER BOX	http://www.conxall.com/Documents/600_700_702
32	Audio Products	Accessories	TE Connectivity AMP Connectors	207313-1	A25074-ND	[0": 0.0]	CONN PLUG HOUSING CPC 4POS BLACK	http://media.digikey.com/pdf/Data_Sheets/Tyco E
33	Audio Products	Accessories	Knowles	2800-253915	2800-253915-ND	[25": 84.46]	SWITCH PUSHBUTTON	http://www.knowles.com/eng/Media/Configuration
34	Audio Products	Accessories	Knowles	2800-253907	2800-253907-ND	[25": 84.46]	SWITCH PUSHBUTTON	http://www.knowles.com/eng/Media/Configuration
35	Audio Products	Accessories	Knowles	009P-253160	423-1228-ND	[1": 24.11, *100": 18.07]	PROGRAMMING ACCESSORIES	http://www.knowles.com/eng/Media/Configuration
36	Audio Products	Accessories	Knowles	2900-253348	423-1236-ND	25 [1": 9.16, *100": 6.872]	SWITCHES	http://www.knowles.com/eng/content/download/5

csv格式数据示例 (digikey.com)

```
{  
  
  "Category1": "Audio Products",  
  
  "Packaging": "",  
  
  "Category2": "Accessories",  
  
  "Description": "GASKET FOR SEALING NEMA APPLIC",  
  
  "Mounting type": "",  
  
  "Lead free": "Contains lead ",  
  
  "Package case": "",  
  
  "Digi-Key Part Number": "458-1048-ND",  
  
  "Brand": "Mallory Sonalert Products Inc.",  
  
  "Manufacturer Part Number": "SG1",  
  
  "Supplier device package": "",  
  
  "Datasheet": "http://www.mallory-sonalert.com/specifications/SG1.pdf",  
  
  "Url": "http://www.digikey.com/product-detail/en/mallory-sonalert-products-inc/SG1/458-1048-ND/528103",  
  
  "Image": "http://media.digikey.com/photos/Mallory Sonalert Products Photos/SG1.jpg",  
  
  "json_filename": "jsons/Audio_Products/Accessories/458-1048-ND.json",  
  
  "Standard package": "1",  
  
  "Attributes": "{\"Category\": \"Audio Products\", \"Product Photos\": \"SG1\", \"Family\": \"Accessories\", \"Color\": \"Black\", \"Series\": \"-\", \"Material\": \"Neoprene\", \"Accessory Type\": \"Gasket, Seal\", \"For Use With/Related Products\": \"Panel Mount Alarms\", \"Other Names\": \"458-1048\"}",  
  
  "Moisture Sensitivity Level": "",  
  
  "Price": "{\"0\": 0.0}",  
  
  "RoHS": " RoHS non-compliant",  
  
  "Stock": ""  
  
}
```



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！



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
