POWERFUL CRO TIPS

Using a **SEO web crawler**

@MrLukeCarthy

#Searchlove

custom **Ecommerce Consultant Luke**Carthy FRESHENER & Founder eCommerce Consultant co. Worked **H**AT **ChemistDirect** with: **RENAULT**

@MrLukeCarthy

#Searchlove



As a SEO, l've failed... (strictly speaking)





Organic traffic is virtually flat..



However, online sales are up by **175%**



More traffic ≠ more sales



More sales ≠ more traffic



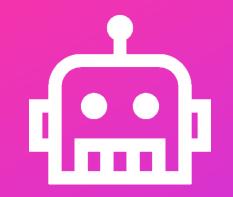












236,000 indexed internal search URLs







Hmmm, we didn't find anything for "supercalifragilisticexpialidocious"

Try a different search term or check out some of our suggestions below.

Purchase History

Are you looking for order status or tracking information? View your Purchase History>

Shop by Category









Desktops

5







Small Appliances

Laptops

Cameras & Camcorders

Customer Service

Visit our customer service>

We're here to help.

Health, Fitness & Sports



Home Audio



Car Electronics & GPS

r

iPod & MP3 Players

Cell Phones



A

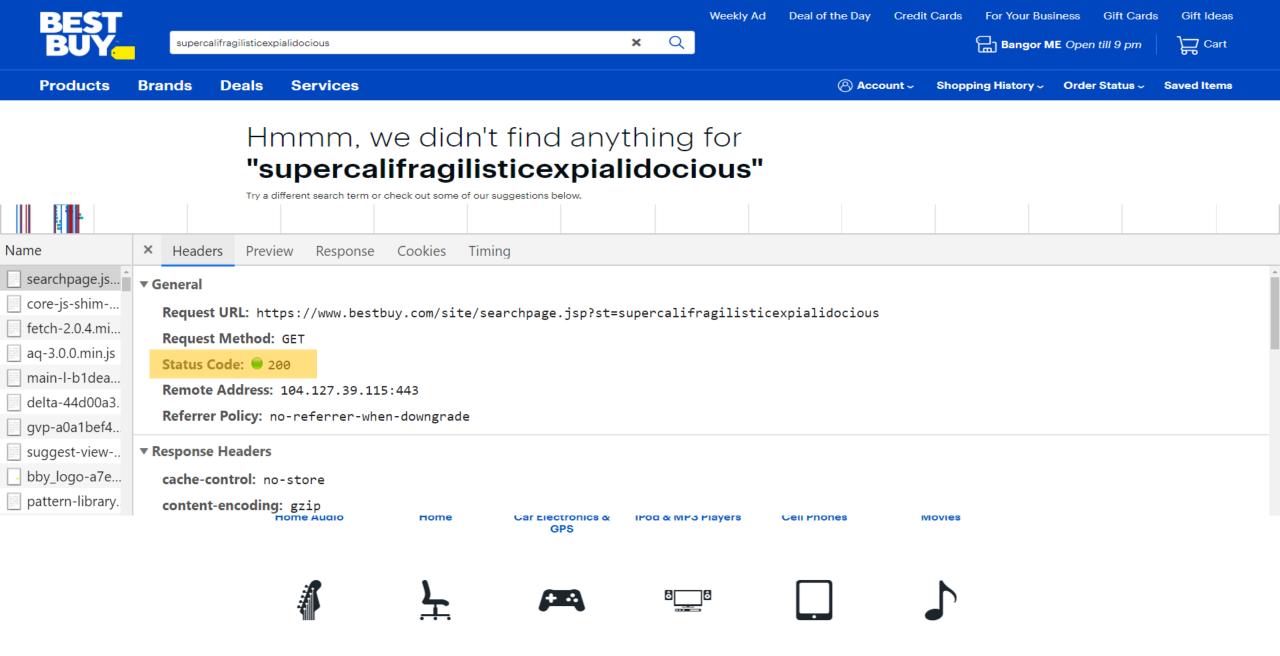












***Zero results** internal searches do not...



STG.

Hello, is it me you're looking for?

GO HOME LIONEL YOU'RE DRUNK

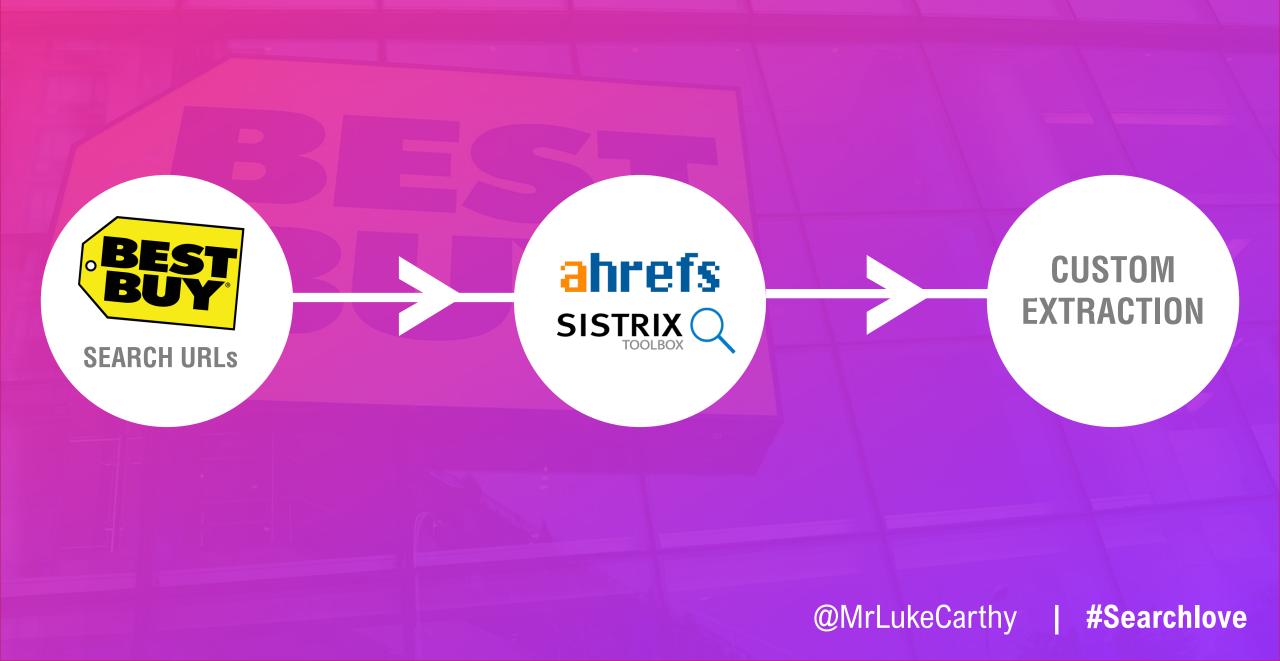
DEEPCRAWL

Sitebulb

Scre@mingfrog







I crawled just a 4% sample of the 236K uRLs...

@MrLukeCarthy

#Searchlove

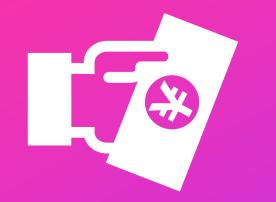
...and surfaced **296** URLs returning 'no results found'

@MrLukeCarthy

#Searchlove

That's **190K** people seeing 'no results found' pages

every month



That's a **HUGE** bunch of missed CRO opportunities





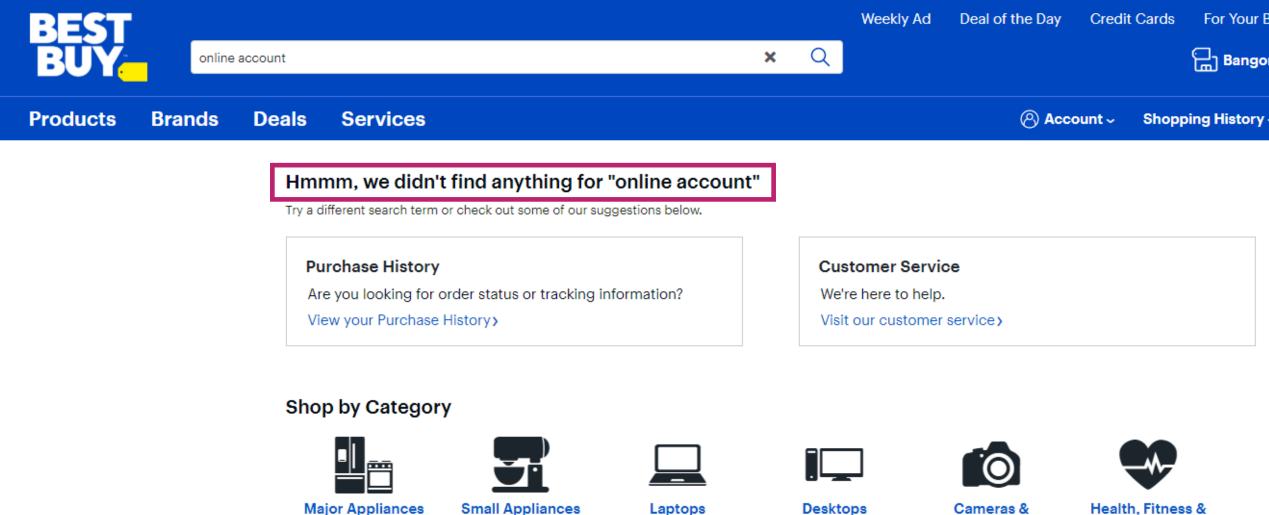
Best Buy done fu\$#&d up...





Well, a little..





Sports



Home Audio

Home

Car Electronics & GPS

ſ

iPod & MP3 Players

,,

0

Cell Phones

Camcorders

....

Movies







How is all this impacting Best Buy?





At least **190K** people aren't finding what they're looking for



Could be **1000S** of potential places where demand is outstripping supply



Not providing a seamless experience for non-commerce search queries



Not selling lines that customers want to buy



30% revenue increase via Internal search optimisation

Ok, now tell us how you got all this cool shit with

CUSTOM EXTRACTION

@MrLukeCarthy

#Searchlove

"Custom extraction allows you to scrape any data from the HTML of web pages when crawling"

https://www.screamingfrog.co.uk/webscraping/





×

Q

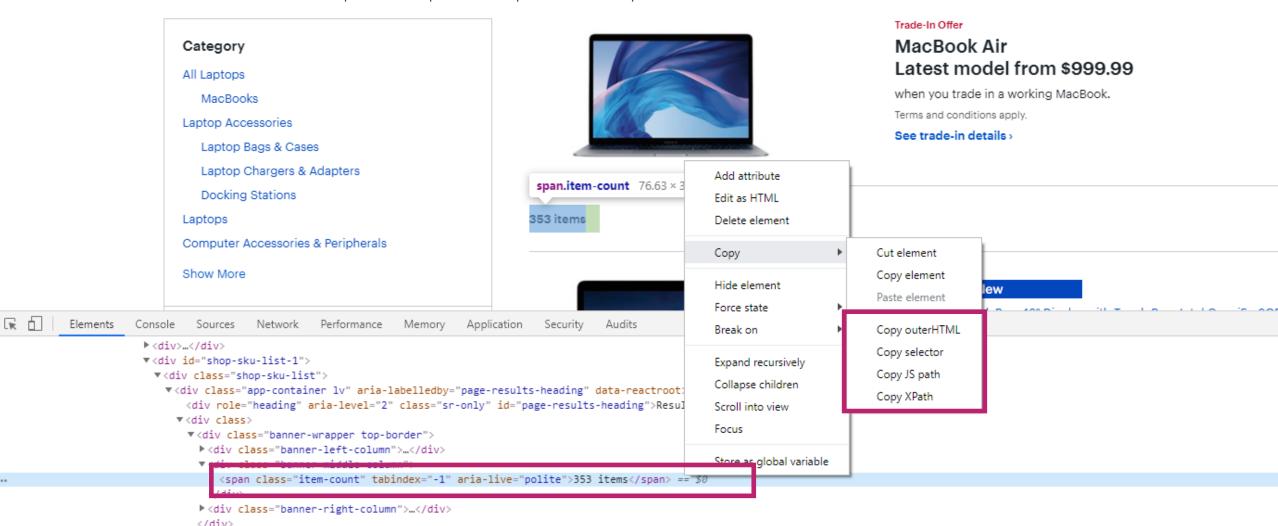
Products Brands Deals Services

macbook

🖉 Account ~ Shopp

Results for "macbook".

Related Searches macbook pro | macbook air | macbook 12" | macbook charger | macbook mini





macbook

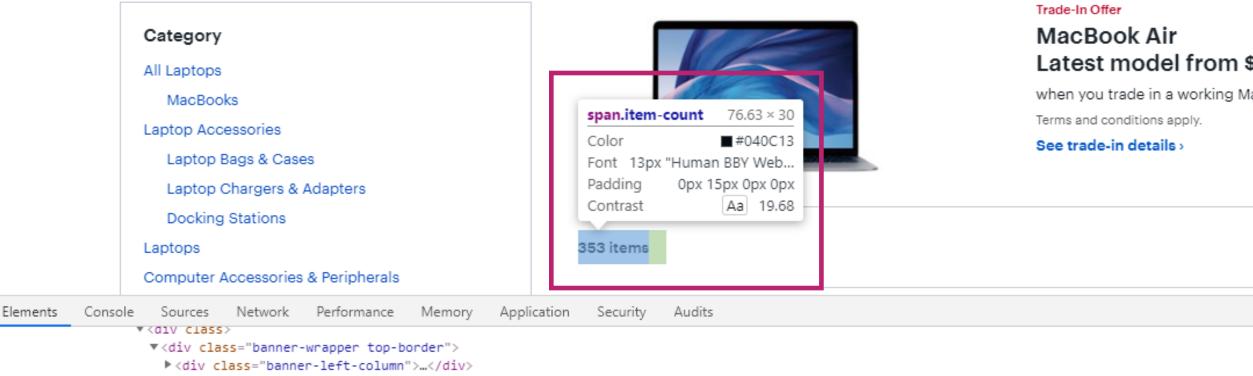
Weekly Ad D

× Q

Products Brands Deals Services

Results for "macbook".

Related Searches macbook pro | macbook air | macbook 12" | macbook charger | macbook mini



V/div_class="bannen_middle.column">

353 items == \$0

>> <div class="banner-right-column">...</div>

(div)

		Weekly Ad Deal of the Day	Credit Cards F	or Your Business Gift Cards Gif	t Ideas
	macbook	× Q		Bangor ME Open till 9 pm	Cart
Brand cbook". es macbook p		<pre>} .shop-sku-list .item-count { padding: > 0 15px 0 0; font-weight: 500; line-height: 30px;</pre>	nt~ Shopping	History ~ Order Status ~ Saved	l Items
ks essories 3ags & Cases Chargers & Ada	ap	<pre>} *, :after, :before { <u>pattern-library.min.css:2</u> webkit box sizing: border box; box-sizing: border-box; }</pre>			
Stations Accessories & I	Pe 	<pre>Inherited from body.size-1 body { pattern-library.min.css:2 font-size: 13px; line height: normal; color: ■#040c13; font-family: Human BBY Web,Arial,Helvetica,sans-serif;</pre>	ntel Core i5	Sort By: Best Match \$1,799.99 Free Apple Music for 3 months (new subscribers only) 8.1 more 0.520.00 year	~
"heading" ari	lv" aria-lab	elledby="page-results-heading" data-reactroot> lass="sr-only" id="page-results-heading">Results		Styles Computed Event Listeners Filter element.style { .shop-sku-list .item-count { padding: b 0 15px 0 0;	OOM Breakpoints Properties >> :hov .cls +_ <u>banner.less:16</u>
ass="banner-m class="item- ass="banner-r	right-column">	> ex="-1" aria-live="polite">353 items == \$0 		<pre>*, :after, :before {</pre>	<u>pattern-library.min.css:2</u>
="list-wrappe ="footer top-	border wrappe	id="main-results" aria-labelledby="main-results-heading">	column span.item-cour	Inherited from body.size-1 body { font-size: 13px; line-height: normal; color: ■#040c13; font-family: Human BBY Web.Aria	<u>pattern-library.min.css:2</u>

Screaming Frog SEO Spider 11.1 - List Mode (Entered Manually)

File	Configuration	Mode Bulk Export	Reports	Sitemaps	visualisations	Crawi Analysis	Licence	нер				
Scr	Spider	roa		C.	Screamin	na Eroa SE() Spide	er 11.1 - List Mode (Entered I	Start	Clear	Crawl 100%	

Scr Spider	rog	🔮 Scre	aming Frog SEO	Spider 1	1.1 - List Mode	(Entered l	Start	Clear	Craw	wl 100%	
Intern	tocol Response Codes URL Pag	File	Configuration	Mada	Bulk Export	Dener	Directives	Hreflang AJA)	K AMP Struc 💌	Overview Site Structure Response Times API	
URL Rewriting	Export	File	Configuration	Mode	BUIK EXPORT	Repon	Search			▼ Summary	
Filter: CDNs		~								Total URLs Encountered: 4978	
Item co 1 item		Scr	Spider	rog	3		nd message 1		+	Total Internal Blocked by robots.txt: 0	
49 item Exclude		50			2				0	Total External Blocked by robots.txt: 0	
3 items Speed		Interes	robots.txt		December 6	adaa 🛛				Total URLs Crawled: 4978	
1155 ite		Intern	URL Rewriting	locol	Response C	odes				Total Internal URLs: 4978	
66 item 8 items HTTP Header		(OKL Kewnung							Total External URLs: 0	
12 item Custom		Filter:	CDNs	🔓 Exp	port		-			▼ SEO Elements	
	Search		1							▼ Internal	
22 item User Interface	Extraction	Item co	Include							All (4978) (100.00%)	
45 item API Access	•	1 item									
148 iter 9 items			Exclude	_						HTML (4976) (99.96%)	
8 items System	•	49 item		_						JavaScript (0) (0.00%)	
17 items		3 items	Speed							CSS (0) (0.00%)	
729 items		1155 it	User-Agent							Images (0) (0.00%)	
6 items		1100 10	User-Agent	_						PDF (0) (0.00%)	
4 items 7 items										Flash (0) (0.00%)	
2457 items		8 items	HTTP Header							Other (2) (0.04%)	
12 items										Unknown (0) (0.00%)	
24 items		12 item	Custom	Sear	rch					▼ External	
108 items		5 items								All (0) (0.00%)	
24 items 11 items		22 item	User Interface	Extra	action					HTML (0) (0.00%)	
39 items			-							IC	
29 items		45 item	API Access	•							Internal
79 items		148 iter									
1 item		9 items	Authentication								
85 items 6 items		9 items	Custom								
(8 items	System						~		
		17 item							Filter Total: 4,978		
1 Export		729 iter									
Name	Value	6 items									
Address	https://www.bestbu								â	HTML	
URL Encoded Address	https://www.bestbu	4 items								Other	
Content Otatus Os da	text/html;charset=U	7 items							U	- Other	
Status Code Status	200 OK										
Indexability	Indexable	2457 it	ems								
Size	342.0 KB										
Title 1	4k 144hz monitor - I	Best Buy									
Title 1 Length	27										
Meta Keyword 1	4k 144hz monitor										

		No results found message 1 +	Total Internal Blocked by robots.txt: 0 Total External Blocked by robots.txt: 0 Total URLs Crawled: 4978 Total Internal URLs: 4978
Custom Extraction			>
		. See our User Guide for examples. The results can be seen in the Custom Tab in the Extraction filter.	
Item count 🗙	CSSPath 🔻	.shop-sku-list .item-count	X Extract Text Attribute (optional)
No results found message X	CSSPath 👻	.shop-zero-search-results .lv .no-results-copy h3.no-results-message	X Extract Text Attribute (optional)
Extractor 3	Inactive 💌		
Extractor 4	Inactive 👻		
Extractor 5	Inactive 👻		
Extractor 6	Inactive		
Extractor 7	Inactive	ges. See our User Guide for examples. The results can be	a cean in the Custom Tab in the Extra
Extractor 8		ges, see our oser ourde for examples. The results can be	e seen in the Custom rab in the Extra
Extractor 9	Inactive	shop-sku-list .item-count	
Extractor 10	Inactive		
		.shop-zero-search-results .lv .no-results-copy h3.nd	Li de la constante de la const
		anop zero searen results av ano results copy risar	o-results-message
		anop zero searen resares av ino resares copy nom	o-results-message
y.com/site/searchpage.jsp?id=		44hz+monitor	o-results-message
y.com/site/searchpage.jsp?id= y.com/site/searchpage.jsp?id= JTF-8		44hz+monitor	_
y.com/site/searchpage.jsp?id=		44hz+monitor	HTML

Good reads on custom extraction

Uproer

https://uproer.com/articles/screaming-frog-custom-extraction-xpath-regex/

LinkedIn Article – Brian Shumway

https://www.linkedin.com/pulse/custom-extraction-screaming-frog-xpath-csspathbrian-shumway

PMG

https://www.pmg.com/blog/how-to-use-xpath-in-screaming-frog/

Search Engine Journal

https://www.searchenginejournal.com/regex-seo-website-data-extraction/234999/

@MrLukeCarthy

#Searchlove



What else can you do with custom extraction?





Extract your product recommendations



"35% of what consumers purchase on Amazon and 75% of what they watch on **Netflix** come from product recommendations." J MIGHT ALSO LIKE

https://www.mckinsey.com/industries/retail/ our-insights/how-retailers-can-keep-up-withconsumers McKinsey &Company



Understand what direct 'bolt on' products customers proactively buy at scale





The first few recs are often the most popular / best selling





Pay close attention to the product price vs the first recs price





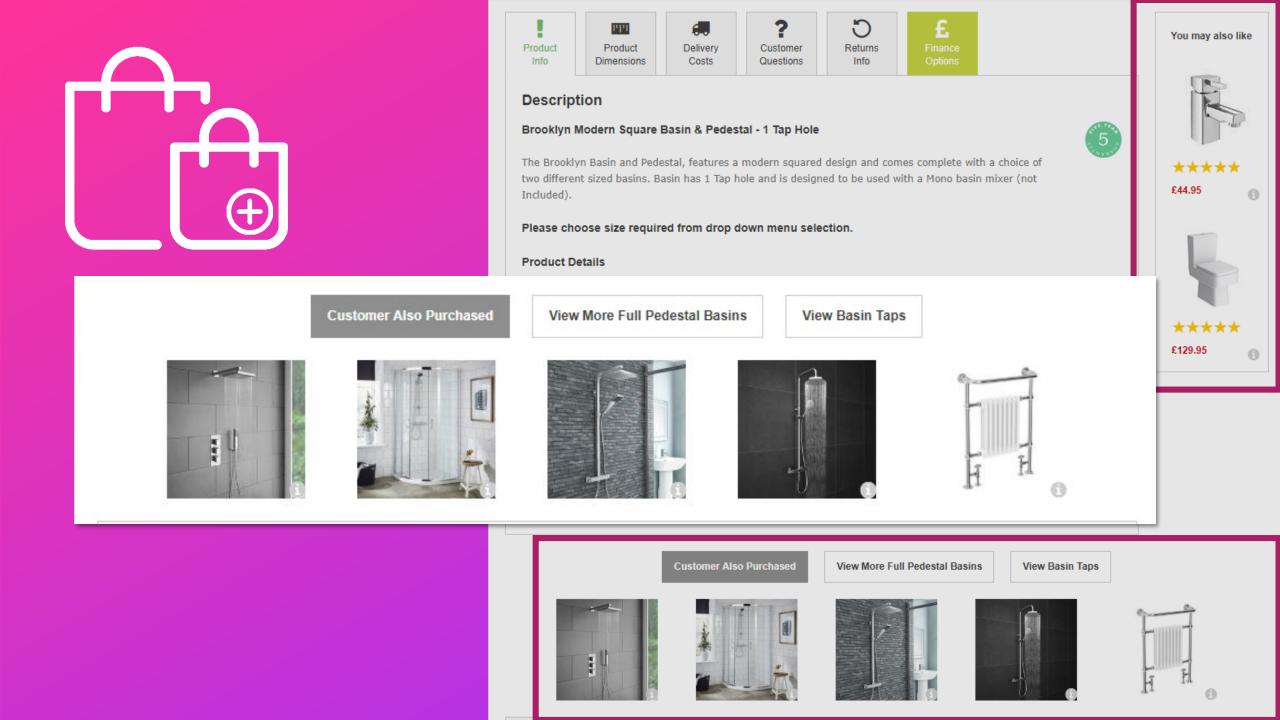
Really powerful when scraping recs from your direct competitors





Even more so when competitors have multi-dimensional recs







36% have no recommendations at all



Were all priced the same or higher

80+%

Appeared to be alternatives rather than complimentary



of out of stock items Don't show recommendations



100% have recommendations showing 56%

Were all priced the same or higher



Appeared to be alternatives rather than complimentary



Appeared to be complimentary rather than alternatives



Shoes = Shoes



Jeans = Shirts



Shirts = Jeans



Shorts = Trainers/Sneakers



Identify your worst / best selling products in every category





Best selling products are often the first items in the category





Competitive ecommerce Intelligence to boost conversion





Identify your competitor's top (and worst) selling products





Work out your competitors search URL pattern



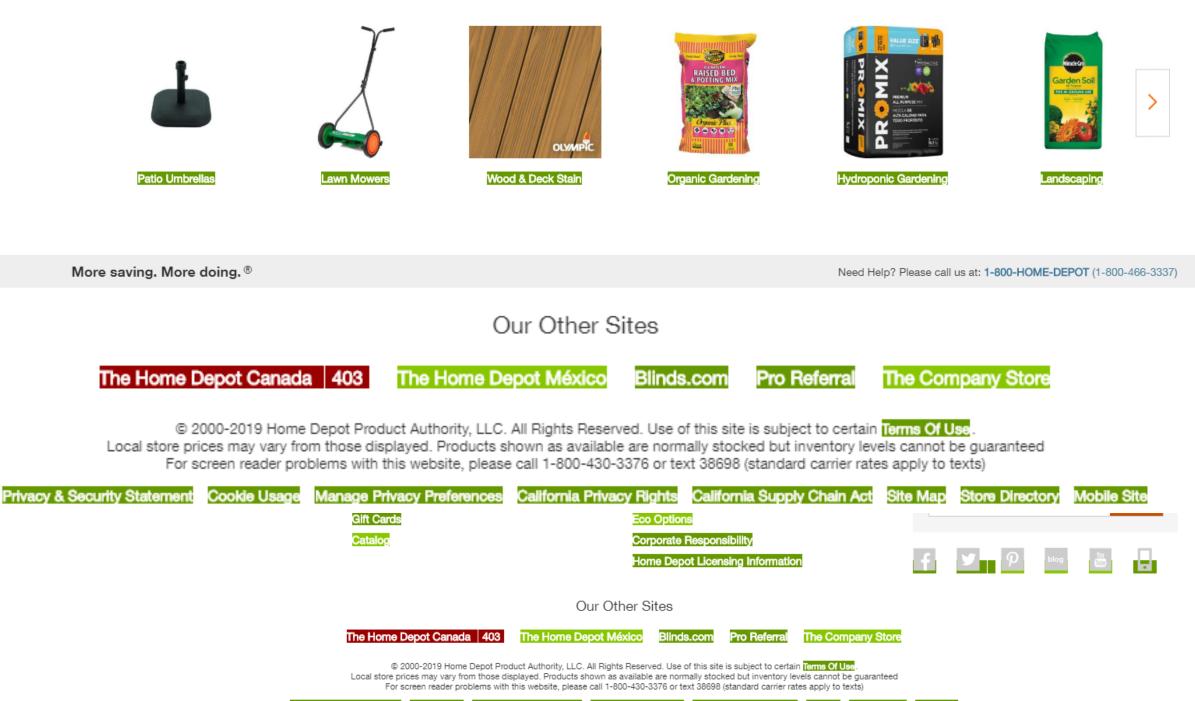
example.com/search?t=[KEYWORD]





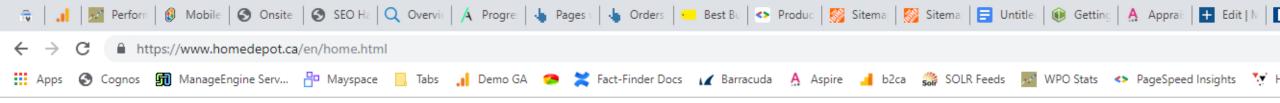
Scrape results for price, reviews and stock and compare to your own





Privacy & Security Statement Cookie Usage Manage Privacy Preferences California Privacy Rights California Supply Chain Act Site Map Store Directory Mobile Site





Access Denied

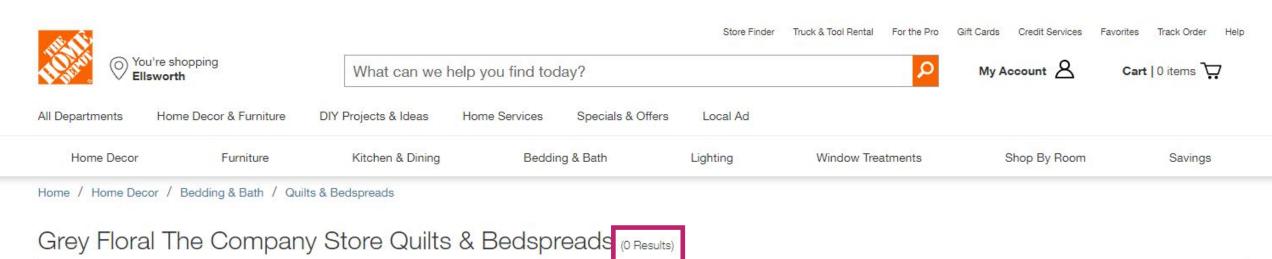
You don't have permission to access "http://www.homedepot.ca/en/home.html" on this server.

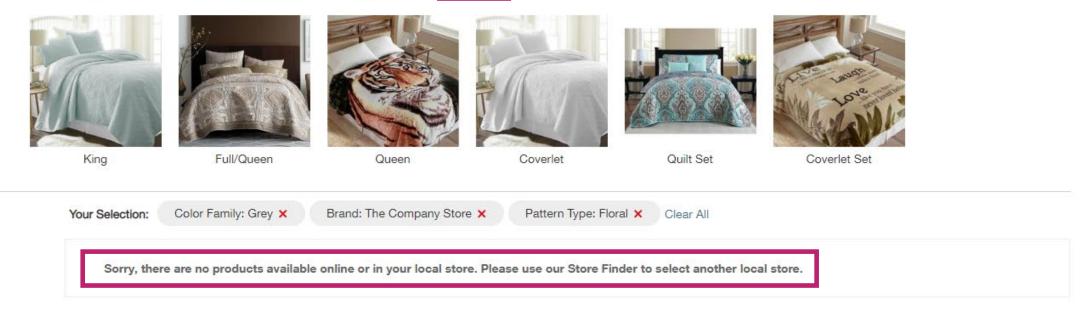
Reference #18.16b77568.1560522504.22cd5932

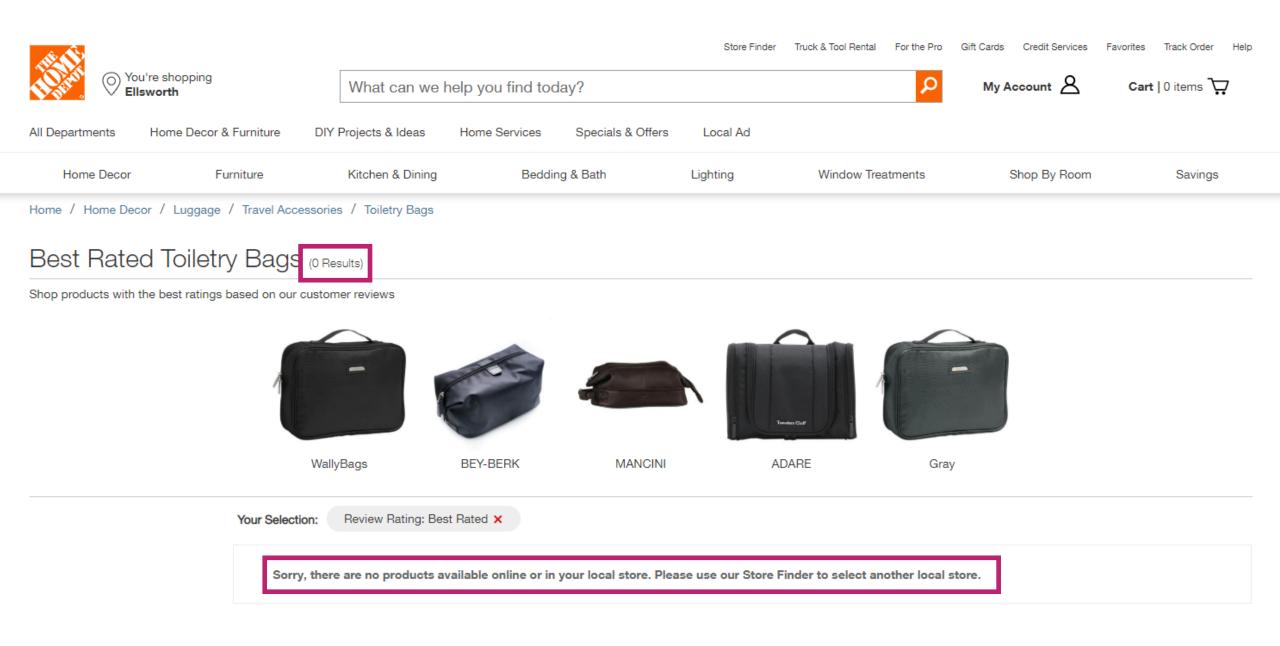


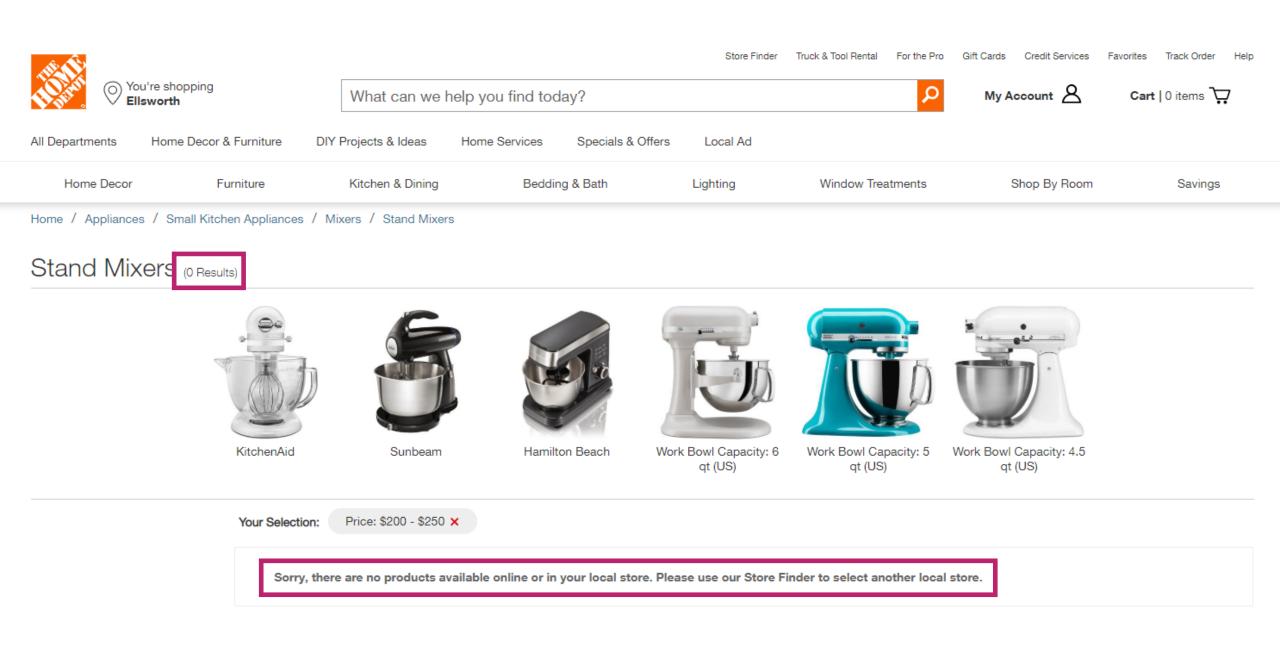
Vet your category navigation URLs in main menus

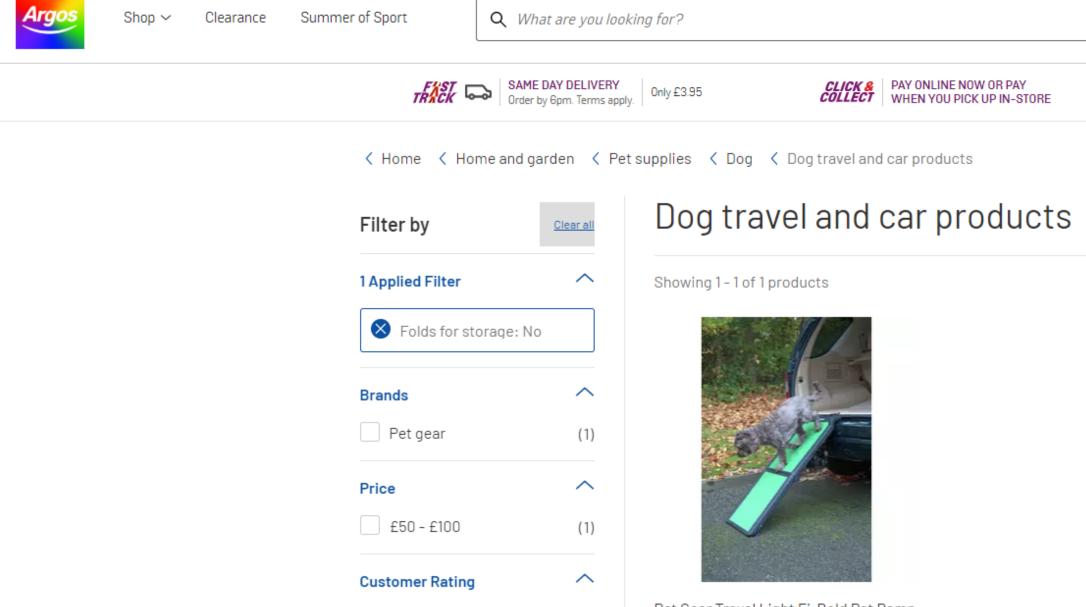












(1)

(1)

Show All

2 2 or more

Pet Gear Travel Light Fi-Bold Pet Ramp

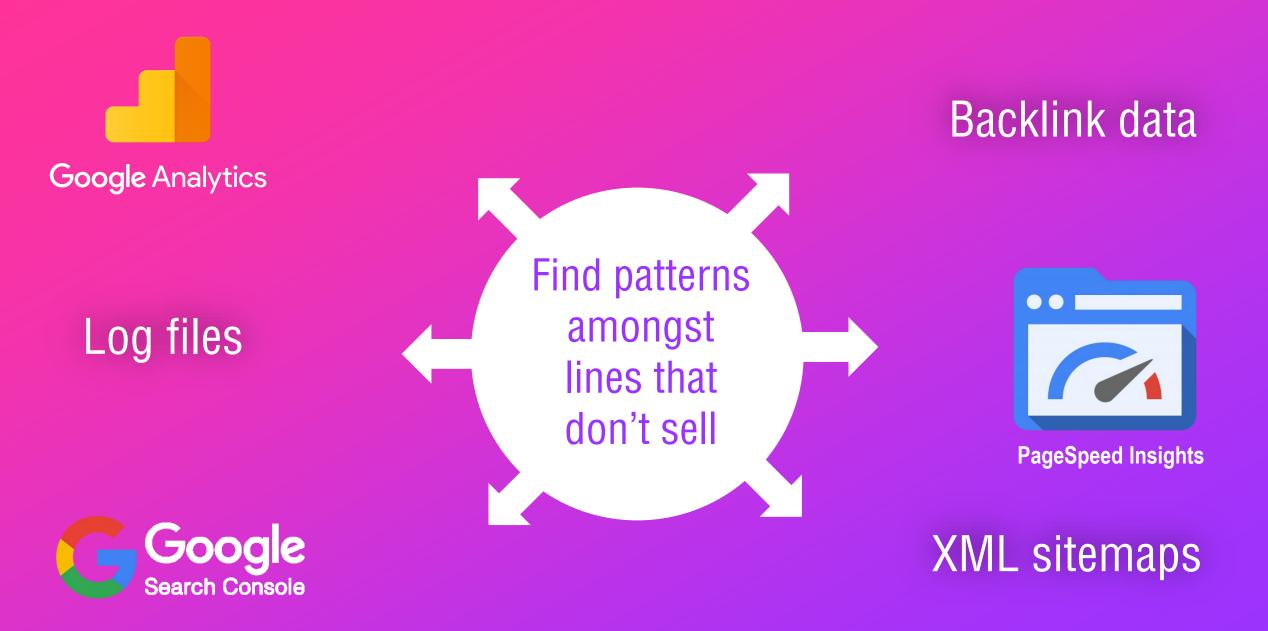
YOU CAN PA WITH THE A

Sort by Rele

SEXY CUSTON

EXTRACTION







It's not just eCommerce...



	Sitebulb	Screaming Frog	DeepCrawl	OnCrawl
Custom extraction	NO (but soon)	YES	YES	YES
Google Analytics	YES	YES	YES	YES
Search Console	YES	YES	YES	YES
Majestic SEO	NO	NO	YES	YES
AHREFS	NO	YES	NO	Coming soon
Logfile Analysis	NO	YES – separate product	YES	Yes – separate product
PageSpeed Insights	NO	Separate crawl	YES (JS injection)	NO

I wrote a thing:

Lukecarthy.com/searchlove

Thank you!



@MrLukeCarthy



/in/lukecarthy



lukecarthy.com