

Spidey ¢ents

Insight Data Engineering, Boston
Dylan Russell

Online Retail Insight



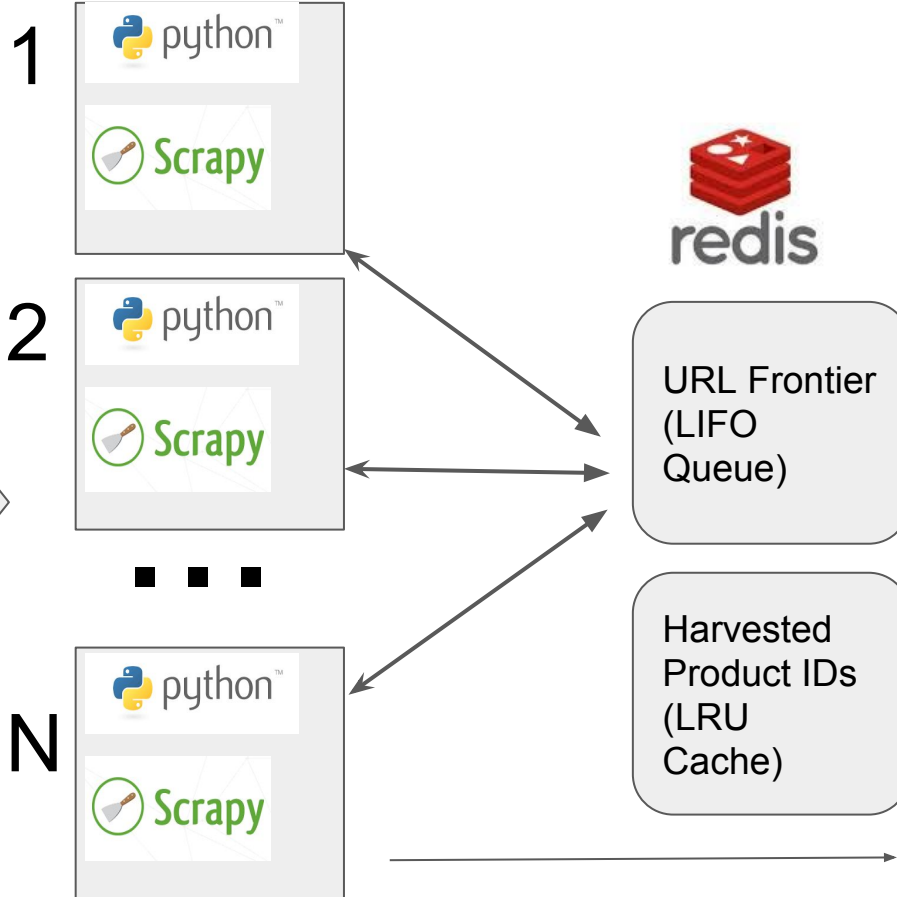
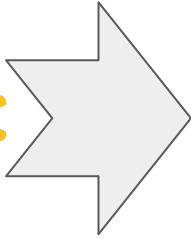
TARGET



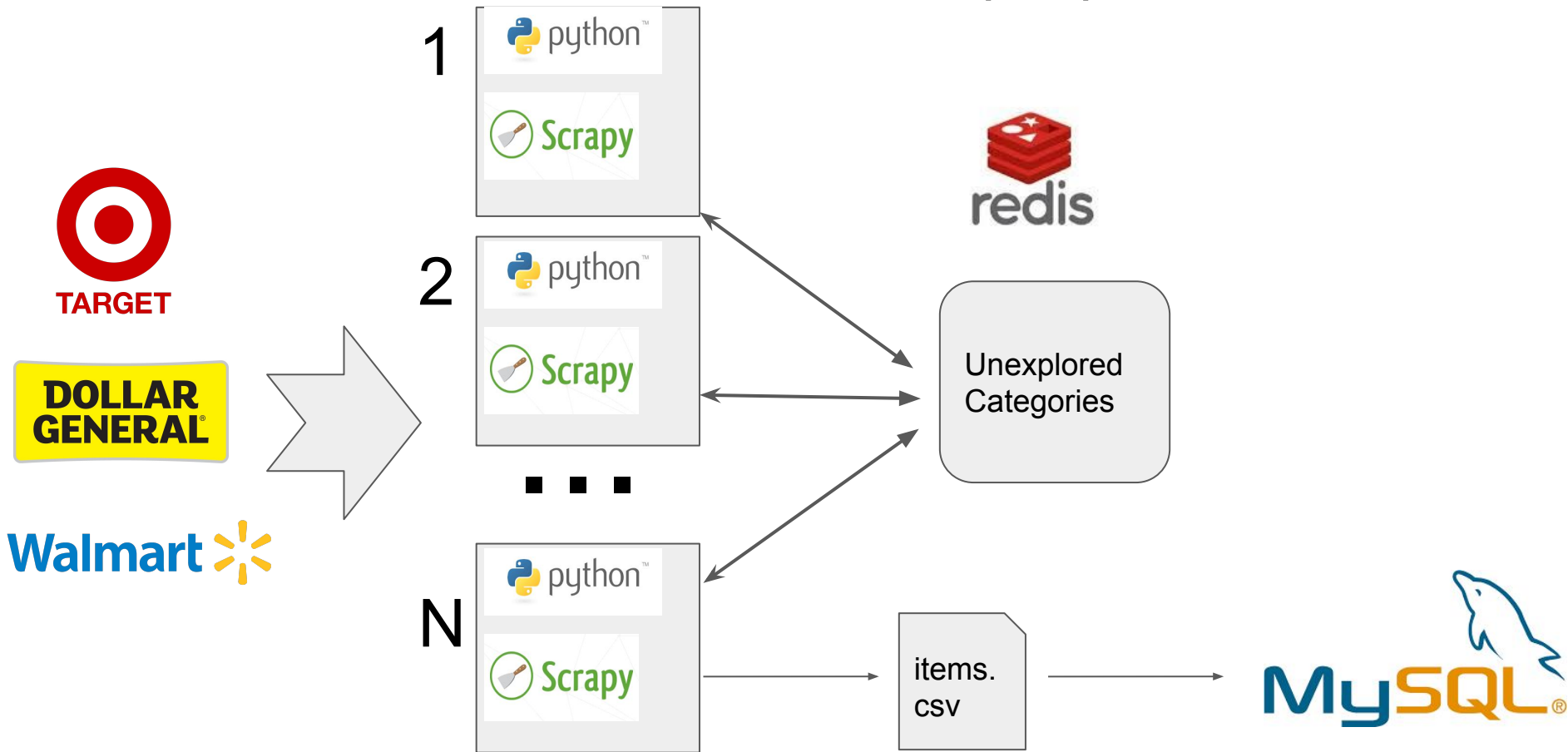
Product Demo

<https://www.youtube.com/watch?v=yj1G3NrjB4A&feature=youtu.be>

Web Crawler



Web Crawler (v2)



Crawling Strategy



Showing 1-40 of 33,958,584 products

Refine by | Price | Top Brands | Store Availability

Sort by | Best Sellers

Departments

- Home
- Books
- Jewelry
- Electronics
- Auto & Tires
- Home Improvement
- Clothing

[See All Departments](#)

Shipping & Pickup

- ☒ Show all
- ☐ 2-day shipping
- ☐ Ship to Home
- ☐ Free Pickup + Discount
- ☐ FREE Pickup
- ☐ FREE Pickup Today

Special Offers

- ☐ Rollback
- ☐ Clearance

BEST SELLER



Apple Watch Series 1 - 38mm - Sport Band - Aluminum ...

★★★★★ 114

\$149.00 List ~~\$249.00~~

2-day shipping
Free pickup today



Pampers Swaddlers Diapers (Choose Size and Count)

★★★★★ 29

\$8.97 - \$67.59

2-day shipping
Free pickup today

[Select sizes](#)

BEST SELLER

BEST SELLER



Sponsored product

Cheez-It Original Baked Snack Crackers 21 oz. Box

★★★★★ 344

\$4.54 22.0¢ / oz

2-day shipping
Free pickup today

BEST SELLER





Blue Box Walmart eGift Card


★★★★★ 50





\$5.00 - \$200.00

How's our new look?

Crawling Strategy




Home ▾ Search 



Showing 1-40 of 4,106 products

Refine by | Price ▾ Top Brands ▾ Store Availability ▾

Sort by | Best Sellers ▾ 

Shipping & Pickup

☐ Show all

☐ 2-day shipping

☐ Ship to Home

☒ Free Pickup + Discount

☐ FREE Pickup

Length

+

Width

+

Brand

—

☒ Trademark Art

See More Brands

...

Height - Top to Bottom

+

Price

—


\$

 to

\$

Go

☒ \$20 - \$50




Trademark Fine Art "Rainbow Seeds Abstract Gold" Can ...

\$37.77 - \$57.53

2-day shipping

Pickup discount eligible




Trademark Fine Art "Bombay Bohemian" Canvas Art by D ...

\$38.67 - \$79.99

2-day shipping

Pickup discount eligible




Trademark Fine Art 'License Plate Map USA' Canvas Ar ...

★★★★★ 1

\$30.28 - \$60.93

2-day shipping

Pickup discount eligible




Trademark Fine Art "Lemon Splash II" Framed Matted A ...

\$21.56 - \$54.58

2-day shipping

Pickup discount eligible








Trademark Fine Art 'Inky Floral III' Canvas Art by S ...

\$38.67 - \$84.92

2-day shipping

Pickup discount eligible



How's our new look?

Crawl Statistics

	Dollar General	Target	Walmart
Product Catalogue Size	~12,000	~1,000,000	~34,000,000
# Items per Crawl	~12,000	~280,000	~16,000,000
# Requests per Crawl	~800	~7,000	~1,500,000
Length of Crawl	~5 minutes	~20 minutes	~6 hours * 6 nodes

Dylan Russell

Soccer

Basketball

Programming

Memorization

