

Identify Best-selling Products on Amazon

Client: A leading eCommerce player in India

Domain: Ecommerce

Offering: Site-specific crawl and extraction

Challenge: Client's requirement was to scrape details of the best-selling products from [Amazon](#) in order to identify the potential best-sellers on their platform. Best seller products from the entire catalog as well as the best sellers under different categories were needed separately. The required data points were Product name, description, specifications, brand, model, , bestseller rank, reviews count and ratings.

```

02 <?xml version="1.0" encoding="utf-8"?>
03 <root>
04 <page>
05 <pageurl>http://www.amazon.com/dp/B007S0LCAC</pageurl>
06 <amazon_info>
07 <product_detail>Deuce MX Grip,Scott USA,219627-1043</product_detail>
08 <url>http://www.amazon.com/dp/B007S0LCAC</url>
09 <price>$12.95</price>
10 <asin>B007S0LCAC</asin>
11 <uniq_id>a0c661bb59736b8fe2bffd6396348016</uniq_id>
12 <product_id>10004</product_id>
13 <shipping_price>$7.18</shipping_price>
14 </amazon_info>
15 </page>
16 <page>
17 <pageurl>http://www.amazon.com/dp/B0050H9XWY</pageurl>
18 <amazon_info>
19 <product_detail>Deuce ATV Grip,Scott USA,217892-1010</product_detail>
20 <url>http://www.amazon.com/dp/B0050H9XWY</url>
21 <price>$13.00</price>
22 <asin>B0050H9XWY</asin>

```

Ecommerce sample data

The Solution: Client provided us with the list of specific data points to be extracted from Amazon. Fresh data sets were needed from Amazon every week. PromptCloud set up web crawlers by programming them in order to fetch the required data fields. This particular case comes under our Site specific crawl offering as the crawler is programmed specifically to fetch data from a particular source website. With our years of expertise in web crawling and scraping, we started delivering clean data to the client soon after the setup was successfully completed. The data delivery format opted for by the client was XML and the data was directly uploaded to client's S3 location.

Benefits:

- Every technically complicated aspect of the crawl and extraction was taken care of by our team
- The setup was complete in a few days and the data was provided in record time
- Client had all the bestselling products from Amazon, category-wise, ready to be fed into their applications. Our advanced tech stack handled millions of records each week with ease
- Monitoring was set up for the source website to detect changes that need modification of our crawlers