





```
In [17]: import requests
from bs4 import BeautifulSoup
import json
import time

def crawl_myntra(base_url, max_pages=2):
    products = []
    page = 1

    headers = {
        'User-Agent': 'Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/91.0.4472.124 Safari/5
    }

    while page <= max_pages:
        url = f"{base_url}&p={page}"
        print(f"Crawling: {url}")

        try:
            response = requests.get(url, headers=headers)
            response.raise_for_status()
            soup = BeautifulSoup(response.text, 'html.parser')

            # Find the script tag containing the product data
            script_tag = soup.find('script', text=lambda t: t and 'window.__myx = ' in t)

            if script_tag:
                # Extract the JSON data
                json_text = script_tag.string.split('window.__myx = ')[1].split('</script>')[0]
                data = json.loads(json_text)

                # Extract product information
                product_list = data.get('searchData', {}).get('results', {}).get('products', [])

                for product in product_list:
                    products.append({
                        'name': product.get('productName', ''),
                        'brand': product.get('brand', ''),
                        'price': product.get('price', ''),
                        'link': f"https://www.myntra.com/{product.get('landingPageUrl', '')}"
                    })

            page += 1

        except requests.RequestException as e:
            print(f"Error crawling {url}: {e}")
            break
```

```

        time.sleep(2) # Be polite, don't overwhelm the server

    return products

if __name__ == "__main__":
    base_url = "https://www.myntra.com/men-tshirts?rawQuery=men-tshirts"
    products = crawl_myntra(base_url)

    print(f"Crawled {len(products)} products:")
    for product in products:
        print(f"Name: {product['name']}")
        print(f"Brand: {product['brand']}")
        print(f"Price: {product['price']}")
        print(f"Link: {product['link']}")
        print("-" * 50)

```

```
script_tag = soup.find('script', text=lambda t: t and 'window.__myx = ' in t)
```

Crawling: <https://www.myntra.com/men-tshirts?rawQuery=men-tshirts&p=2> (<https://www.myntra.com/men-tshirts?rawQuery=men-tshirts&p=2>)

Crawled 100 products:

Name: Urbano Fashion Men Teal Green Slim Fit Tropical Printed Pure Cotton T-shirt

Brand: Urbano Fashion

Price: 307

Link: <https://www.myntra.com/tshirts/urbano+fashion/urbano-fashion-men-teal-green-slim-fit-tropical-printed-pure-cotton-t-shirt/12377258/buy> (<https://www.myntra.com/tshirts/urbano+fashion/urbano-fashion-men-teal-green-slim-fit-tropical-printed-pure-cotton-t-shirt/12377258/buy>)

-----

Name: The Roadster Lifestyle Co. Pure Cotton T-shirt

Brand: Roadster

Price: 199

Link: <https://www.myntra.com/tshirts/roadster/the-roadster-lifestyle-co-pure-cotton-t-shirt/1996777/buy> (<https://www.myntra.com/tshirts/roadster/the-roadster-lifestyle-co-pure-cotton-t-shirt/1996777/buy>)

-----

Name: BULLMER Tribal Printed Oversized Cotton T-shirt

Brand: BULLMER

In [ ]:

