

# Jewellery Sites Extraction

-Actowiz Team

-Date: 2025/02/18



ACTOWIZ<sup>TM</sup>  
Appliance of Data & Insights



## Steps Explained:

### 1. Install Dependencies:

- Run **pip install -r requirements.txt** to install all required Python modules.

### 2. Configure env File:

- Add your Scrape.do proxy key to the .env file.

### 3. Add Category URLs:

- Place the category URLs you want to scrape into the appropriate `input_files/*_categories.json` file.

### 4. Run the Scraper:

- Execute the Command **python run.py <PLATFORM> <REGION>** (e.g., `python run.py chanel us`).

### 5. Crawler Running:

- The system will: Read the input category files and start extraction process.

### 6. Scrape product listing pages & Gather Product Details

### 7. Save structured data in JSON and Excel formats.

- The scraped data will be saved in the `output_files/` directory

### 9. End:

